National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

Characterization of the *fem-2* gene of *Caenorhabditis elegans*

BY

Angela Christine McGregor

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

DEPARTMENT OF GENETICS

Edmonton, Alberta

Fall, 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-06507-3

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Angela McGregor

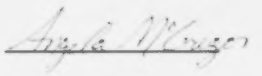
TITLE OF THESIS: Characterization of the *fem-2* gene of *Caenorhabditis elegans*

DEGREE: Masters of Science

YEAR THIS DEGREE GRANTED: 1995

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided neither in the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Angela McGregor

11118-81 Avenue,

Edmonton, AB, T6G 0S5

Date: August 24, 1995

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Characterization of the *fem-2* gene of *Caenorhabditis elegans*" submitted by Angela McGregor in partial fulfillment for the degree of Master of Science.

Dr. David Pilgrim

Dr. John Bell

Dr. Ross Hodgetts

Dr. Ellen Shibuya

Date: May 17, 1995

Abstract

The regulation of sex determination in *Caenorhabditis elegans* involves a complex set of genetic interactions. In this study, the *fem-2* gene was cloned based on its molecular position in the *C. elegans* genome. A genomic clone from this region was used to screen a cDNA library. The cDNA clones were sequenced and aligned with the genomic sequence to determine intron/exon boundaries and the largest open reading frame. This sequence correlates with a single transcript of 1.8 kb. A database search determined that the FEM-2 predicted protein shows sequence similarity to a number of mammalian and yeast protein phosphatases, and suggested that FEM-2 protein may have phosphatase activity. Several additional experiments allowed me to continue both a genetic and a molecular analysis of *fem-2* . This research project has provided crucial information about the *fem-2* gene product and its role in sex determination.

Acknowledgments

Many thanks to the Pilgrim lab. I would like to thank Dave Pilgrim, my supervisor, for his help, guidance and especially his patience with me over the years. I am grateful for his optimistic outlook on research that often kept me going when experiments were not working. Thanks to the grad students in the lab, Dave Hansen, Morris Maduro and Petra Jackle-Baldwin. They made life in the lab interesting and lively. In addition, thanks to Ramesh Bhambhani for help maintaining our worm stocks and for numerous lunch hours of cards. I would like to give a special thanks to Troy Johnson who started off as a lab-mate and ended up being a lifelong friend and now husband-to-be. He helped me tremendously with my work and always was very helpful. Finally, thanks to some of my friends who put up with my three years of no social interaction and to my family who always provided their support.

Table of contents

	Page
Introduction	1
General biology	1
Sex-determination	4
fem-2 phenotype	14
Physical approach to the cloning of *fem-2*	17
Brief statement of goals	18
Materials and Methods	
Strains and growth conditions	20
Escherichia coli strains and culture conditions	20
Microinjection of nematodes	20
Isolation of *fem-2* cDNA clones from library	21
DNA sequencing	22
BLAST search	22
Single worm PCR	23
Heat shock of nematodes	23
EMS mutagenesis	24
Primer list	24
Tc1 insertion screen	25
Results	
cDNA isolation and sequencing	27
Analysis of the DNA sequence	31
Minigene construction and rescue	32
Heat shock-minigene construct	35
fem-2 suppressor screen	38
Tc1 insertion screen	40

Discussion

Isolation and characterization of cDNA clones..... 42

Heat shock studies..... 48

Sequence analysis..... 50

Genetic screens..... 59

Summary..... 63

Bibliography..... 110

Appendix..... 119

List of Tables

Table Page

1. *C. elegans* strains..... 66
2. Recombinant plasmids..... 68

List of Figures

Figure		Page
1.	Genes involved in sex determination in *C. elegans*	71
2.	Molecular model of sex determination	73
3.	Minimum rescuing region for *fem-2*	75
4.	DNA sequence of *fem-2*	77
5.	Construction of cDNA sequence	82
6.	*fem-2* cDNA/genomic arrangement	84
7.	*fem-2* homologue alignment	86
8.	Sequence comparison (conserved motifs I-VI)	88
9.	Conserved motifs alignment	90
10.	*fem-2* minigene insert (AM001+ and AM001-)	92
11.	Rescue of *e2105* XO phenotype	94
12.	PCR amplification of transgenic worms	96
13.	Heat shock results	98
14.	Suppressor screen	100
15.	Tc1 insertion event	102
16.	Intronic repeat	104
17.	Model for FEM-2 activity	106
18.	5' ends of the cDNA clones	108

Abbreviations

ABA	abscisic acid
BLAST	Basic Local Alignment Search Tool
bp	base pair
cDNA	complementary DNA
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dITP	2'-deoxyinosine 5'-triphosphate
DNA	deoxyribonucleic acid
dTTP	2'-deoxythymidine 5'-triphosphate
EMS	ethylmethanesulfonate
kbp	kilobase pairs
mRNA	messenger RNA
ORF	open reading frame
PCR	polymerase chain reaction
RACE	Rapid Amplification of cDNA Ends
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
TBE	Tris-borate, ethylene-diaminetetra-acetic acid
UTR	untranslated region

INTRODUCTION

During development, the generation of sexual dimorphism is one of the most universal processes in the animal kingdom. Sexual determination, and the underlying genetic and molecular mechanisms that direct it, has been studied in great detail in both the fruit fly and the nematode. In the nematode *Caenorhabditis elegans*, understanding the events that lead to the different sexual phenotypes is crucial to identifying the complete set of interactions among genes in the developmental pathway. The *fem* genes (for feminization) play an important role during development and act at a branch point in the sex determination pathway (Hodgkin, 1986). *fem -2*, along with *fem -1* and *fem -3*, are responsible for spermatogenesis in the germ line of hermaphrodites and males, as well as for male somatic development (see Figure 1). The mapping and characterization of *fem-2* cDNA clones and the sequencing of the gene were essential steps towards understanding the genetic and molecular role of FEM-2 in determination of sex in *C. elegans*.

General Biology

Caenorhabditis elegans is well suited to genetic studies for several reasons (Brenner, 1974; Wood, 1988). It is small, easy to grow and hermaphrodites are able to reproduce by self-fertilization or by cross-fertilization using males. It has a large range of differentiated cells that are common to most metazoans and due to the transparency of the animal, every cell in the developing nematode can be studied under the light microscope. This has made it possible to use differential interference contrast (DIC) or Nomarski microscopy (Nomarski, 1955) to show that the entire cell lineage from egg to adult, in living worms is essentially invariant (Sulston *et al.*, 1983). Mutant worm stocks can be maintained indefinitely and inexpensively by freezing in liquid nitrogen,

which means that many parental lines can be preserved for up to 20 years with no changes in genetic background. The complete worm life cycle takes about three days at 20°C under optimal conditions (Brenner, 1974). The developing worm proceeds through four larval stages (L1, L2, L3 and L4) which are punctuated by molts. A new cuticle is synthesized and the old cuticle is shed (Wood, 1988). Sperm is produced in the hermaphrodite gonad during the fourth larval stage then switches to oogenesis at the L4 molt. The two sexes, male and hermaphrodite, grow to an adult size of about one mm in length (Brenner, 1974). Hermaphrodites have a total of 959 somatic nuclei, while males have 1031 (Sulston *et al.*, 1983). Natural populations consist of mostly hermaphrodites which are self-fertile because they are capable of producing sperm for a short time during development. Males arise through the spontaneous loss of one X chromosome and are therefore infrequent in natural populations (Hodgkin *et al.*, 1979). In the laboratory, males can be crossed with the hermaphrodites to produce broods of half males and half hermaphrodites (Brenner, 1974).

Another advantage of this organism for developmental studies is the fact that it has a very small haploid genome size, about 100, 000 kbp (Sulston and Brenner, 1974), with genes being relatively small, introns small and few, as well as containing few repeated sequences (Emmons, 1988). The small genome size has made it possible to begin to construct a complete physical map of the genome, which is presently greater than 95% complete (Coulson *et al.*, 1988; Coulson *et al.*, 1991) and to sequence the entire worm genome (Wilson *et al.*, 1994). To date, Wilson *et al.* (1994) have sequenced the majority of chromosome III, the central region of chromosome II, and a small number of cosmids near the center of the X chromosome (Sulston *et al.*, 1992; Wilson *et al.*, 1994). There are six linkage groups or chromosomes in *C. elegans*.

How complex patterns of cells form within initially undifferentiated groups of cells is a fundamental problem in development. *C. elegans* is an ideal candidate for studying such processes as temporal control of development (Kenyon, 1985), programmed cell death (Hengartner and Horvitz, 1994), neuronal guidance (Siddiqui, 1990) and sex determination (Kuwabara and Kimble, 1992), to name a few. Gradually, it seems possible that some parts of the development of this animal will become completely understood with respect to the genes and proteins that are involved. Developing cells not only have to decide what type of cell they will become but also, whether they are male-specific or female-specific. Therefore, sex determination is an appealing system because the cell must choose between only two alternative states rather than a large number of fates such as epidermal cells, muscle cells, neurons, etc. The sum of these individual developmental decisions results in the somatic differences between the male and the hermaphrodite. The commitment to a single sexual fate and the maintenance of that decision involves a number of key sex determining genes (reviewed in Kuwabara and Kimble, 1992; Villeneuve and Meyer, 1990). The study of how sexual fates are specified in *C. elegans* will help to define regulatory networks that affect other aspects of development.

The control of sexual dimorphism in *C. elegans* involves a cascade of major regulatory genes that direct development along either male or female pathways. In order to piece together how these genes regulate sex determination, it is essential to understand the sex-specific events that occur during the development of the hermaphrodite and of the male. Sexual development in the worm involves a large number of well characterized events that affect almost every tissue, most which result from different patterns of cell division in the two sexes. Embryonic development is almost identical in the two sexes, however post-embryonic cell lineages are highly divergent (Hodgkin *et al.*, 1989). The result of these lineages is 30% to 40% of all somatic cells and all germ cells are sex-

specific (Wood, 1988). The first embryonic sexually dimorphic event occurs at approximately 13 hours after fertilization. At this time, four cephalic neurons (CEM) in the hermaphrodite undergo cell death as well as two hermaphrodite specific motorneurons (HSN) in the male (Sulston and Horvitz, 1977). Therefore, at the end of embryogenesis the hermaphodite has 556 nuclei and the male has 558. In the post-embryonic lineages a relatively small number of precursor cells undergo sex-specific divisions that result in 281 hermaphrodite-specific nuclei and 351 male-specific nuclei (Hodgkin *et al.*, 1989). Only 122 nuclei are common to both sexes. These later dimorphic features begin to occur shortly after hatching in the L1 animal. The sexually dimorphic cells make up a large number of tissues including the germ line, the intestine, the somatic gonad, the vulva, the sex muscles and associated neurons and male tail. The other events that will be addressed later include the onset of spermatogenesis and oogenesis. In the hermaphrodite spermatogenesis begins at the L3/L4 molt and after approximately eight hours the germ line switches to the production of oocytes. In the male spermatogenesis begins during the L3 stage and continues throughout the adult stage of the worm. There are a large number of details involving sexual dimorphism that will not be discussed here because they are not relevant but have been reviewed in Hodgkin *et al.* (1989), Hirsh *et al.* (1976) and Hodgkin (1988).

Sex Determination

The study of sex determination in *C. elegans* provides insight into different types of cell signaling pathways that control developmental choices of cells and organisms. The regulation of sex determination involves a complex set of genetic interactions. In *C. elegans* the first step in the sex determination pathway is controlled by the sex chromosome (X) to autosome (A) ratio (Madl and Herman, 1979). This ratio ultimately determines whether the organism will develop as a male (XO:AA) or a hermaphrodite (XX:AA) (Madl and Herman, 1979). The X/A ratio acts as a signal that is directed at a

small number of major sex determining genes (Hodgkin and Brenner, 1977; reviewed in Kuwabara and Kimble, 1992). This signal controls the state of these genes, which then promotes development along a specified sexual pathway. The growing number of genes that respond to this signal act to direct three independent processes: dosage compensation, germ line sexual development, and somatic sexual development (reviewed in Villeneuve and Meyer, 1990). How the X/A ratio is read by a cell is unknown, but the study of the genes that respond to this signal may lead to the understanding of the whole process of sexual differentiation and why it is organized in this way.

The nematode, like many other organisms, possesses a specific mechanism to compensate for a difference in sex linked gene dosage (Hodgkin, 1983). X-specific transcript levels must be equalized between males (XO) and hermaphrodites (XX) by either up-regulating the genes on the single X chromosome in males or by down-regulating one of the two X chromosomes in the hermaphrodite. Meneely (1990) has proposed a model that dosage compensation is achieved by 'turning down' the level of X-linked gene expression in XX animals to that of XO animals. The genes that control both dosage compensation and sex determination are *xol-1*, *sdc-1*,*-2* and *-3* (Miller *et al.*, 1988; Villeneuve and Meyer, 1990; Nusbaum and Meyer, 1989; DeLong *et al.*, 1993) (see Figure 1) and were identified as X-linked recessive mutations. These genes presumably play a role in either reading the X/A ratio and/or transmitting the information to pathways controlling sex determination or dosage compensation. The X-linked gene *xol-1* (for XO lethal) is required for the male mode of sex determination and dosage compensation. Loss-of-function mutations in this gene cause embryonic or early L1 lethality in XO animals, but have no effect on XX animals (Miller *et al.*, 1988; Meneely, 1990). Of the small number of XO eggs that do hatch and grow to L1, most show some transformation toward the hermaphrodite fate (Miller *et al.*, 1988). In contrast, *sdc-1* and *sdc-2* (for sex determination and dosage compensation) mutations in XX animals cause

masculinization and increased levels of X-linked gene expression while XO animals are unaffected (DeLong *et al.*, 1993). These two genes are therefore important for promoting hermaphrodite development. The null phenotypes of loss-of-function mutations in *sdc-1* and *sdc-2* are somewhat different. All *sdc-1* mutations do not completely transform all XX animals toward the male fate; some develop as normal hermaphrodites. Unlike *sdc-2* mutations, *sdc-1* does not cause extensive XX-specific lethality despite increased levels of X-linked gene expression. In addition, all *sdc-1* mutations exhibit a strong maternal rescue such that *sdc-1* XX progeny from *sdc-1* heterozygous mothers are normal hermaphrodites (Villeneuve and Meyer, 1990). In contrast to *sdc-1* and *sdc-2*, *sdc-3* mutations cause disruptions in dosage compensation but do not cause any obvious sexual transformation (DeLong *et al.*, 1993). The *sdc-3* gene exhibits unusual genetic behavior in that its functions in sex determination and dosage compensation are independent. The analysis of genetic interactions between these genes has demonstrated that the feminization caused by *xol-1* mutations (in surviving animals) is completely suppressed by mutations in either *sdc-1* or *sdc-2*, causing XO animals to be viable wild type males. These results have suggested that *xol-1* is the first gene in a regulatory hierarchy that controls the male/hermaphrodite decision (Miller *et al.*, 1988; Nusbaum and Meyer, 1989). Figure 1 describes the pathway that indicates where these genes are located in the hierarchy. In this model, *xol-1* responds to the X/A ratio and acts to down regulate the *sdc* genes in XO animals. The *sdc* genes are placed upstream of *her-1* in this pathway because a mutation in the *her-1* gene can completely suppress the masculinization caused by *sdc* mutations (Nusbaum and Meyer, 1989; reviewed in Meneely, 1990). These three genes are clearly important for proper functioning of the sex determination and dosage compensation pathways, however they do not function alone, but act at an early step that is common to both. The genes that branch off at this point to direct dosage compensation (but have no effect on sex determination) in XX animals are referred to as the 'dosage compensation dumpies'

(*dpy-21*, *dpy-26*, *dpy-27*, and *dpy-28*) (Plenefisch *et al.*, 1989; Meneely, 1990). Mutations in these autosomal genes are deleterious or lethal to XX animals but not to XO animals. They also result in an XX-specific dumpy phenotype (Hodgkin, 1983). Since this portion of the sex determination and dosage compensation pathways is not directly involved with the *fem* genes, the description of the genes involved is not detailed.

There are a number of genes that act to control sex determination but not dosage compensation. These genes are involved in determining the sexual fate of cells in all tissues that exhibit sexual dimorphism. How these genes function and interact appears to be different in the soma and the germ line, therefore they will be discussed separately. The seven major genes that direct somatic sexual phenotype have been grouped into three classes: *her-1*, three *fem* genes and three *tra* genes (for reviews see Hodgkin, 1990; Villeneuve and Meyer, 1990; Kuwabara and Kimble, 1992). *her-1* (for hermaphrodite) loss-of-function mutations cause XO animals to be transformed into hermaphrodites. The three *tra* genes (for transformer) exhibit the same mutant phenotype which is masculinization of XX animals. The three *fem* genes also have the same mutant phenotype which all cause feminization of XO and XX animals Somatic sex determination refers to the development of all sex-specific tissues except the germ line. The genes that are involved act in a cascade of negative regulatory interactions that ultimately set the activity state of *tra-1*, which is both necessary and sufficient to direct female somatic development (Hodgkin and Brenner, 1977; Hodgkin, 1980). According to the regulatory pathway proposed by Hodgkin (Hodgkin, 1980; Doniach and Hodgkin, 1984), when the X/A ratio is high, *tra-1* is active and a hermaphrodite soma results, when the X/A ratio is low, *tra-1* is inactive and a male soma develops. Briefly, the *her-1* gene is required for male but not hermaphrodite development (Hodgkin, 1980; Schedin *et al.*, 1994). Loss-of-function mutations in this gene result in the complete transformation of XO animals into normal fertile hermaphrodites while XX animals are unaffected. Two

semidominant gain-of-function *her-1* mutations cause variable degrees of masculinization of XX animals but not complete sexual transformation (Trent *et al.,* 1988; Perry *et al.,* 1993; Schedin *et al.,* 1994). *her-1* acts to repress the activity of the *tra-2* and *tra-3* genes, which are next in the regulatory hierarchy (Figure 1). The *tra-2* gene activity is required for hermaphrodite sexual development and is dispensable in males (Klass *et al.,* 1976; Hodgkin and Brenner, 1977). Loss-of-function *tra-2* mutations cause masculinization of XX animals, with the strongest alleles transforming XX animals into non-mating pseudomales which never produce oocytes (Hodgkin and Brenner, 1977). A number of dominant gain-of-function alleles of *tra-2* cause feminization of the XX germ line producing spermless females but have no apparent effects on XO animals (Doniach, 1986; Schedl and Kimble, 1988). One *tra-2* allele that causes a mixed phenotype not only feminizes the germ line of XX animals but also causes partial feminization of XO germ line and soma (Doniach, 1986; Goodwin *et al.,* 1993). The *tra-3* gene is also required for hermaphrodite development and like *tra-2*, loss-of-function mutations transform XX animals into non-mating pseudomales (Hodgkin and Brenner, 1977; Hodgkin, 1980). In contrast to *tra-2* however, *tra-3* mutations have a strong maternal rescue meaning that *tra-3/tra-3* XX progeny from *tra-3* heterozygous mothers develop as normal hermaphrodites (Hodgkin, 1980). Zygotic rescue can also eliminate the masculinization effect, in that *tra-3* heterozygous progeny from a *tra-3/tra-3* mother and a wild type male are also normal hermaphrodites.

Next in the proposed epistasis pathway are the *fem* genes, which *tra-2* and *tra-3* together act to repress (Nelson *et al.,* 1978; Kimble *et al.,* 1984; Doniach and Hodgkin, 1984; Hodgkin, 1986). Loss-of-function mutations in any one of these three genes causes both XX and XO animals to develop as spermless females. While all three *fem* genes are necessary for spermatogenesis in XX and XO animals and for male somatic development, there are significant differences between them. Most *fem-1* loss-of-

function mutations cause complete feminization of XX animals and complete, or almost complete, feminization of XO animals (Doniach and Hodgkin, 1984). *fem-1* (null) mutations transform both XX and XO animals into fertile females, but XO animals are completely feminized only when both mother and zygote are homozygous *fem-1* (null) (Doniach and Hodgkin, 1984). In addition, of the XX *fem-1/fem-1* progeny from a heterozygous *fem-1* mother only about 20% are self fertile. Alleles of *fem-3* also cause some degree of feminization of XX and XO animals, but the maternal effect is different from that of *fem-1* and *fem-2* (Hodgkin, 1986). XX animals require two wild type copies of *fem-3* in order to be normal self fertile hermaphrodites, regardless of the maternal contribution. This implies haplo-insufficiency that cannot be rescued by maternal product (Hodgkin, 1986). Alternatively, in XO animals *fem-3* only feminizes if both mother and zygote are homozygous *fem-3*. From these observations Hodgkin (1986) suggested that the three *fem* products are required in different amounts with *fem-3* product being required in the largest amount and *fem-2* product in the smallest amount. Gain-of-function *fem-3* mutations have been found that cause masculinization of the hermaphrodite germ line and suppress the feminizing effect of *fem-1* or *fem-2ts* mutations (Barton *et al.*, 1987). This suppression does not affect the feminization of the male soma that is seen in most *fem-1* and *fem-2* XO mutants. The *fem* genes are responsible for down regulating *tra-1*, which is both sufficient and necessary to direct female somatic development (Hodgkin and Brenner, 1977; Hodgkin, 1986). Most loss-of-function *tra-1* mutations lead to masculinization of XX animals, while gain-of-function mutations cause the feminization of XO and XX animals (Hodgkin, 1980; Hodgkin, 1983). The activity states of these seven genes in the cascade direct the somatic sexual phenotype of the worm. This pathway predicts that in the XX animal *her-1* activity is low, *tra-2* and *tra-3* activity states are high, in turn the activity of the three *fem* genes is low and ultimately, *tra-1* activity is high which results in a female soma. In XO animals *her-1* activity is high and this leads to low activity of *tra-2* and *tra-3*, high

activity of the *fem* genes and low activity of *tra-1* resulting in a male soma (Figure 1) (Hodgkin, 1980; Doniach and Hodgkin, 1984). A current molecular model for how these genes interact is shown in Figure 2 which illustrates a signal transduction pathway initiated by the *her-1* gene product. Detailed discussion of this model and the genes that are involved will follow.

All of the above mentioned genes, as well as a number of additional genes, are involved in germ line sex determination, but their functions and interactions are different from that of somatic sex determination. Germ line sexual differentiation dictates the type of gamete that will be produced by the germ line. Similar to somatic sex determination, this pathway involves a cascade of negative interactions. A significant difference lies in the role of the *fem* genes. These three genes together activate spermatogenesis in both males and hermaphrodites, as well as promote male somatic development by repressing *tra-1* activity (Hodgkin, 1986). *tra-1* also appears to play a minor role in determination of germ line sex, that is, strong loss-of-function mutations can cause reduced spermatogenesis, and can sometimes induce oogenesis in both XO and XX animals (Schedl *et al.*, 1989). This suggests that some level of *tra-1* activity is necessary for proper spermatogenesis, a function that is opposite to its somatic role. Another aspect of germ line sexual differentiation that is unique is the temporal modulation that must occur in the hermaphrodite germ line that enables it to produce sperm for a short time then switch to production of oocytes. Normally, the hermaphrodite ovotestis begins to make sperm at the L3/L4 molt, then switches to the production of oocytes which continues through adulthood (Hirsh *et al.*, 1976). The average number of sperm produced by a wild type hermaphrodite is about 320 (Wood, 1988). This germ line modulation is achieved by a number of genes that are germ line specific including *fog-1,-2*, and *-3*, and the six *mog* genes (Barton and Kimble, 1990; Schedl and Kimble, 1988; Ellis and Kimble, 1995; Graham and Kimble, 1993; Graham *et al.*, 1993). The *fog-1* gene (for

feminization of the germ line) is required for specification of germ cells to become sperm and acts with the three *fem* genes at the end of the sex determination pathway to direct spermatogenesis in males and hermaphrodites (Barton and Kimble, 1990). Loss-of-function mutations in this gene cause XX animals to become fertile females and feminizes XO animals, eliminating the production of sperm but having no effect on the somatic cells in either sex. It appears that the *fog-1* gene is required for the onset of spermatogenesis during early development in both XX and XO animals and for the continuation of spermatogenesis in males (Barton and Kimble, 1990). In comparison, *fog-2* acts together with the *fem* genes and *fog-1* in specifying the production of sperm (Schedl and Kimble, 1988). Specifically, *fog-2* is required for the onset of hermaphrodite spermatogenesis but is not required in XO males. A loss-of-function *fog-2* mutation causes the feminization of the XX germ line, similar to *fog-1*, however, XO males are unaffected (Schedl and Kimble, 1988). It has been proposed that *fog-2* may regulate the *fem* genes and *fog-1* to activate them in the hermaphrodite germ line to promote spermatogenesis by negatively regulating *tra-2*. The third *fog* gene, *fog-3*, is required for germ cells to differentiate into sperm instead of oocytes (Ellis and Kimble, 1995). Loss-of-function mutations in this gene cause XX and XO animals to produce oocytes only. Using an epistasis analysis, Ellis and Kimble (1995) have shown that *fog-3* acts at the end of the sex determination pathway and is essential for germ cells to differentiate as sperm. Another germ line specific gene is *mog-1* (for masculinization of germ line) (Graham and Kimble, 1993). In the absence of *mog-1* activity, spermatogenesis does not switch to oogenesis in the hermaphrodite germ line, therefore germ cells that would normally have become oocytes are transformed into sperm (Graham and Kimble, 1993). These XX animals show no somatic masculinization. XO *mog-1* animals are normal males which exhibit normal mating behavior and produce functional sperm. This suggests that *mog-1* is necessary for the switch from spermatogenesis to oogenesis in the hermaphrodite (Graham and Kimble, 1993). It has

been demonstrated that maternal *mog-1* is necessary for proper embryogenesis and most progeny that do not receive any *mog-1* product from their mother die as embryos (Graham and Kimble, 1993). The purpose of maternal *mog-1* in embryogenesis is not understood. Mutations in the other five *mog* genes have the same effect on germ line sex determination, that is in hermaphrodites spermatogenesis continues beyond the point at which oogenesis is to begin and consequently, no oocytes are produced (Graham *et al.*, 1993). *mog-2* - *mog-6* mutants also exhibit maternal effect lethality. By analyzing *mog* mutations in combination with other sex determination mutants, Graham *et al.* (1993) have suggested that these genes act before *fem-3* in the sex determination pathway and may negatively regulate *fem-3*. The above mentioned genes are important in the specification and maintenance of germ line sexual fate. The genetic control of germ line sex determination must be tightly regulated in order for the proper switch from spermatogenesis to oogenesis to take place in hermaphrodites.

Although many of the sex determination genes are genetically well characterized, and some have been cloned and analyzed at the molecular level, it is not completely understood how the genes interact with one another. A model is emerging that involves intercellular signaling and transcriptional regulation. The generation of a signal that mediates cell-cell interactions may regulate a signal transduction pathway that leads to the transcriptional activation of one of the late acting sex determination genes. A speculative model for the molecular regulation of somatic sexual fate has been described by Kuwabara and Kimble (1992) and by Kuwabara *et al.* (1992) (Figure 2). In this model *her-1*, which is thought to promote male development (Hodgkin, 1980), acts as a negative regulator of *tra-2*. The *her-1* gene is predicted to encode an extracellular secreted protein of low molecular weight (Perry *et al.*, 1993). *tra-2* encodes a predicted membrane protein which may act as a receptor for HER-1 and transduce its signal to the cytoplasm of the cell (Kuwabara *et al.*, 1992). In this case, and all others that follow,

her-1 refers to the genetic locus and HER-1 to the protein product of that locus. It has been proposed that TRA-2 is inactivated due to the binding of HER-1, consequently, male development ensues. *tra-3* has been cloned and is predicted to encode a protein with sequence similarity to a calpain (Tom Barnes, personal communication), a calcium-regulated, cytosolic protease found in various species. The predicted TRA-3 protein may act by cleaving TRA-2 in order to activate it. The proteins encoded by the *fem* genes are predicted to be intracellular and at least *fem-3* may interact with the intracellular domain of the TRA-2 protein (Spence *et al.*, 1990; Ahringer *et al.*, 1992; Spence and Kuwabara, personal communication). Because each *fem* gene is essential for male somatic development and spermatogenesis, it has been suggested that the three protein products interact with one another as a multisubunit complex to promote spermatogenesis (Kuwabara and Kimble, 1992).

The *fem-1* gene encodes a protein that contains six ankyrin repeats (Spence *et al.*, 1990) which have been implicated in mediating protein-protein interactions (Lux *et al.*, 1990). This suggests that FEM-1 may interact with either TRA-2 directly or interact with the other *fem* proteins to bind TRA-2. *fem-3* encodes a novel intracellular protein and contains a short sequence in its 3' untranslated region (UTR) that is responsible for its negative regulation (Barton *et al.*, 1987; Ahringer and Kimble, 1991). The final gene in the pathway, *tra-1*, is alternatively spliced to produce two proteins with zinc-finger motifs (Zarkower and Hodgkin, 1992). It has been suggested that TRA-1 may act as a transcription factor that may regulate downstream genes that direct female differentiation (Zarkower and Hodgkin, 1992; 1993). The molecular characterization of the genes in the *C. elegans* somatic sex determination pathway has been used to predict a novel type of signal transduction system that would direct somatic sex determination, as shown in Figure 2 (based on Kuwabara *et al.*, 1992). In this speculative model the genetic regulatory pathway and the molecular identity of each of the genes is combined to

illustrate how sexual fate is determined in XX and XO animals. In XO males, HER-1 is active and binds to TRA-2 and prevents its interaction with the FEM proteins. In turn, the FEM proteins are then able to interfere with *tra-1* activity and male development occurs. In contrast, in XX hermaphrodites, TRA-2 is available to bind to one or more of the FEM proteins, thus allowing TRA-1 to enter the nucleus and transcriptionally activate downstream genes that promote female development (Kuwabara *et al.*, 1992). Although this model seems the most appealing presently, as more information is obtained about the activities of the proteins involved, alternate mechanisms may be proposed.

fem-2 Phenotype

Genetic epistasis evidence predicts that *fem-2* acts at a branch point in the control of germ line and somatic development. Homozygous *fem-2* mutant alleles result in feminization of the XO and XX somas, as well as the germ line, affecting two aspects of sex determination. *fem-2(e2105)* is the strongest allele of *fem-2* and is the most likely candidate for a null allele (Hodgkin, 1986). This allele of *fem-2* causes complete transformation of XO animals into fertile females only at 25°C, yet complete germ line feminization of XX animals occurs at 25°C and 20°C (Hodgkin, 1986). XO animals at 20°C are intersex with incomplete male tails and hermaphrodite-like somatic gonads that often burst open at a partially formed vulva. XO animals at 25°C are fertile females with low fertility (viable egg production) compared to feminized XX animals (Hodgkin, 1986). Another allele of *fem-2, b245*, is a temperature sensitive allele which only exhibits the mutant phenotype at 25°C implying that this allele may have some residual *fem-2* function (Kimble *et al.*, 1984; Hodgkin, 1986). At 25°C, XO *b245* animals have reduced sperm production, induction of oogenesis and formation of a vulva in a small percentage of animals. In XX animals raised at 25°C, sperm production is eliminated (Kimble *et al.*, 1984). This mutant phenotype implies that the wild type function of *fem-2* is to promote the male developmental program. Genetic evidence suggests that it

functions to positively regulate spermatogenesis as well as negatively regulate *tra-1*, a gene that promotes female somatic development (Hodgkin, 1986). The *fem-2* gene product or mRNA, despite its role in masculinization, is provided by the maternal germ line and demonstrates a classic maternal effect. This means that the *fem-2* phenotype is only exhibited when a homozygous *fem-2* mutant zygote does not inherit any wild type *fem-2* product from its mother. If the mother is heterozygous for a *fem-2* mutation, her progeny will inherit some *fem-2*(+) product and will develop as hermaphrodites. This maternally inherited *fem-2*(+) product is sufficient to cause production of some sperm but not the wild type amount (Hodgkin, 1986). This demonstrates that partial masculinization is caused by the maternal contribution of *fem-2*(+) and that *fem-2* activity may possibly be regulated post-transcriptionally. A similar maternal effect is seen for *fem-1* and *fem-3* (Doniach and Hodgkin, 1984; Hodgkin, 1986). How a masculinizing factor could be provided by a female tissue has a number of possible explanations (Hodgkin, 1986). It may be that the *fem-2* product does not have an effect on the germ line sexual phenotype and some other influence acts extrinsically to make the choice between sperm and oocytes. Another possibility is that *fem-2* product is made too late in the development of the female gonad to affect the gamete type. Why there is any maternal effect at all (and only with the *fem* genes) is not fully understood.

Using the strongest *fem-2* allele, *e2105*, Hodgkin (1986) was able to determine the interactions of *fem-2* with other sex determining genes and found that they were almost identical to those observed for *fem-1* (Doniach and Hodgkin, 1984). The *fem-2* phenotype is completely epistatic to *tra-2* and *tra-3* mutations, meaning the masculinizing effects of these mutations can be suppressed by *fem-2*. By constructing the double mutants and observing their phenotypes, Hodgkin (1986) demonstrated that *fem-2;tra-2* and *fem-2;tra-3* XX and XO animals are females. This epistasis is complete at 25°C for the *fem-2;tra-2* strain and at both 20°C and 25°C for the *fem-2;tra-3* double mutant. In

contrast, *tra-1* mutations can suppress the somatic feminization of the *fem-2* mutation but, the germ line effect cannot be completely suppressed (Hodgkin, 1986). At 20°C, *fem-2 tra-1* XX animals have a male soma and produce some functional sperm, while XO animals are normal males. At 25°C, *fem-2 tra-1* XX animals also have a male soma, but their germ line produces only a few oocyte-like cells and no sperm (Hodgkin, 1986). This suppression at 20°C demonstrates that *fem-2* is not absolutely required for spermatogenesis to occur, unlike *fem-1* and *fem-3* which appear to be necessary at all temperatures for spermatogenesis to occur normally (Hodgkin, 1986; Barton *et al.*, 1987; Doniach and Hodgkin, 1984). All three *fem* genes share a common characteristic in that they all show a strong maternal effect. This suggests that coordinated *fem* activity may be controlled post-transcriptionally (Doniach and Hodgkin, 1984). Animals that have weak temperature sensitive mutations in more than one *fem* gene exhibit a more extreme feminization than that of any single *fem* temperature sensitive mutants (Kimble *et al.*, 1984). More complex interactions between *fem-2* and *fem-1* and *fem-3* have been examined and it has been demonstrated that a *fem-3* gain-of-function (gf) mutation can suppress the self-sterility of temperature sensitive alleles of either *fem-1* or *fem-2* (Barton *et al.*, 1987). However, this suppression does not occur when stronger *fem-1* or *fem-2* mutant alleles are examined which indicates that the *fem-3 (gf)* mutation cannot compensate for a complete lack of *fem-1* or *fem-2*.

As mentioned previously, the phenotypes and maternal effects observed in mutants of each of the *fem* genes are similar. The temperature sensitive period (TSP) is also similar for these three genes, suggesting that they may act together during sex determination (Nelson *et al.*, 1978; Kimble *et al.*, 1984; Hodgkin, 1986). The analysis of temperature sensitive periods involves carrying out temperature shift experiments with temperature sensitive mutations. Temperature sensitive mutant alleles are predicted to produce wild type protein product at the permissive temperature and when shifted to the

restrictive temperature the protein product become non-functional. This non-functional protein could be due to a number of things such as improper folding of the protein or the destruction of binding ability of the protein. The TSP for a gene may reflect the time the gene product is synthesized or when its activity is required or a combination of the two. In the case of *fem-2(b245)*, staged worms were shifted from permissive temperature to restrictive temperature, or *vice versa* in a series of reciprocal shift experiments (Kimble *et al.* 1984). The *fem-2* TSP in the hermaphrodite germ line begins 10 hours post-hatch (first larval stage or L1) and ends 18 hours post-hatch (mid L2). In contrast, the *fem-2* TSP for the male germ line is much longer, starting in L2 and continuing through adulthood. This TSP for the male corresponds to the period of time before spermatogenesis begins in the male germ line. The hermaphrodite germ line TSP precedes the brief period of sperm production that occurs during the L4 stage (Hirsh *et al.*, 1976). The TSP for the *fem-2* effect on the male somatic gonad is different than that of the germ line. It extends from hatching to 12 hours post-hatch. This time period correlates to the first time when signs of sexual dimorphism occur in the embryo (Kimble and Hirsh, 1979). The TSPs for the three *fem* genes in the germ line are overlapping and precede the time period when sperm are produced (Nelson *et al.*, 1978; Hodgkin, 1986; Kimble *et al.*, 1984). The TSPs that have been determined for the other genes in the sex determination pathway are consistent with the various landmarks during the embryonic and larval development that distinguish males from hermaphrodites (Klass *et al.*, 1976; Kimble *et al.*, 1984; Hodgkin, 1986, Nelson *et al.*, 1978). The temporal order of the TSPs for these genes is entirely consistent with the order of gene function inferred from genetic epistasis (Villeneuve and Meyer, 1990).

Physical approach to the cloning of fem-2

An alignment of physical and genetic maps for the left end of Linkage Group III suggested a location for *fem-2* within the genome (Pilgrim, 1993). This analysis placed

the *fem-2* mutant alleles between two Restriction Fragment Length Polymorphisms (RFLP), *eP95* and *eP64*, which represented 0.2 map units or 50 kbp. Cloned DNA fragments from this region were tested for their ability to rescue the *fem-2* mutant phenotype, following germ line transformation (Pilgrim *et al.*, in press) (Figure 3). Rescue was confirmed by the restoration of self-fertility to homozygous *fem-2(b245)* transgenic animals at the non-permissive temperature. Three overlapping cosmid clones and one lambda clone were each able to restore spermatogenesis in transgenic animals (Pilgrim *et al.*, in press). The rescuing region was subsequently narrowed down to 3.0 kbp using plasmid subclones of the cosmid and lambda rescuing clones. The boundaries of the rescuing region are defined by the ends of the subclones DP#DBP026 and DP#DBP141 (Table 2). These studies determined the minimal region sufficient to rescue the XX *fem-2* mutant phenotype and were an essential step towards the cloning of the *fem-2* gene.

Brief statement of goals

The overall goal of this project was to determine what type of protein *fem-2* encodes, through sequence analysis, and to characterize how *fem-2* might function in *C. elegans* sex determination. This analysis would clarify the role of *fem-2* in the proposed signal transduction pathway and test the genetic prediction that *fem-2* acts as both a positive regulator of spermatogenesis and a negative regulator of female somatic and germ line development. Results from this project may also suggest how the *fem-2* protein interacts with the other identified proteins in the pathway. Both a genetic and a molecular approach were used to reach these goals. The first specific aim of this research was to clone and sequence the *fem-2* gene. The next goal was to use a genetic approach to search for genes that might interact with *fem-2*. This was particularly aimed at isolating mutations that suppress the feminization effect of the *fem-2* mutation, if any exist. Another goal of this project was to identify a null mutant of *fem-2* by screening for

a transposable element insertion into the *fem-2* coding region. A clear understanding of the null phenotype of *fem-2* is essential to the construction of models for the function of *fem-2*. A complete lack of *fem-2* product may result in a completely novel phenotype that cannot be predicted from other weak alleles of *fem-2*. In addition, the identification of a true genetic and molecular null is very important to the interpretation of the gene's role. This research project was focused on the molecular characterization of the *fem-2* gene and has provided essential information to the advancement of our knowledge of how *fem-2* might function in sex determination. This analysis has also led to many new questions about how *fem-2* may fit into a molecular model. Genetic and molecular analysis has shown that the *fem-2* gene plays a central role in sex determination and is likely to act as an important regulator of sexual fate throughout development. By studying a simple model organism, these experiments may help to understand other complicated developmental processes that involve similar signal transduction pathways.

Materials and Methods

Strains and growth conditions

Nematode stocks were maintained and handled as described by Sulston and Hodgkin (1988). The strains were maintained at 20°C, except where otherwise indicated. See Table 1 for complete list of strains used. The isolation and characterization of the *fem-2* alleles used in this work are described by Kimble *et al.* (1984) and Hodgkin (1986). For most purposes, XO animals were isolated from strains containing a *him-8* (high incidence of males) mutation (Hodgkin *et al.*, 1979)

Escherichia coli strains and culture conditions

The *E.coli* strain used in this research was XL1-Blue as a host strain for the cDNA library. This strain was also used to maintain plasmid subclones of the lambda ZapII vector by *in vivo* excision as described by the manufacturer (Stratagene). XL1-Blue genotype: supE44, hsdR17(r_k^-,m_k^+), recA1, endA1, gyrA96, thi-1, relA1, lac^-, F' [proAB+ $lacI^q$ lacZΔM15 Tn10(tet^r)] (Stratagene)

Cultures of XL1-Blue were grown in liquid or solid (1.5% agar) L-broth media containing the appropriate antibiotics as described in Sambrook *et al.* (1989). For long term storage of bacterial strains, glycerol stocks were prepared as described by Sambrook *et al.* (1989). Sterile 50% glycerol was added to a final concentration of 15% to cells freshly grown to saturation.

Microinjection of nematodes

Young adult hermaphrodites, homozygous for *fem-2(b245ts)*, were raised at the permissive temperature (20°C) and transformed with cloned DNA plasmids using the microinjection technique as described by Mello *et al.* (1991). The worms were injected

with a mixture of the plasmid pRF4 which contains the dominant *rol-6(su1006dm)* marker and the desired plasmid (Mello *et al.* 1991; Kramer *et al.* 1990). If the worm is successfully transformed, the pRF4 plasmid causes the animal to move in a distinct rolli g fashion. Following injection, the hermaphrodites are shifted to the restrictive temperature (25°C) and allowed to self-fertilize. The progeny from this hermaphrodite are examined for the *rol-6* phenotype and for self-fertility. With all transgenic lines, a stable transmitting line of worms was obtained at 20°C, by selecting rolling animals, before testing for *fem-2* rescue. Typically, 50 to 100 worms were injected to obtain one to four transgenic lines.

Isolation of *fem-2* cDNA clones from library

A mixed stage cDNA library (prepared from XX animals of the wild type N2 strain) (Barstead and Waterston, 1989) in the lambda ZAPII/Eco RI vector (Stratagene) was a generous gift from B. Barstead (Washington University, St. Louis, Missouri). The procedure used to plate and screen the library was described in a manual provided by Stratagene Cloning Systems. To screen the library, phage were plated on five large 150 mm x 15 mm LB plates with XL1-Blue strain bacteria and incubated at 37 °C for six to eight hours. The plates were refrigerated for two hours at 4 °C then transferred onto 137 mm Hybond-N (Amersham) nylon filters. Plaque lifts were performed as described in Sambrook *et al.* (1989). Filters were then probed with a 15 kbp *Xba*I fragment (see Figure 3) containing the region of DNA that is able to rescue the *fem-2* mutation. The DNA probe was radiolabelled using the T7 Quickprime kit (Pharmacia) as described by the manufacturer. From approximately 10^5 recombinant phage screened, 43 positive clones were isolated. Plaques of interest were taken from the agar plate with a pasteur pipette and stored as described in Sambrook *et al.* (1989). Each phage stock was titered and the cloned inserts contained within the lambda were recovered to form a phagemid, pBluescript SK(-), using the process of *in vivo* excision as described in the Stratagene

manual. The lambda genome contains the phagemid pBluescript, which in turn contains the cDNA inserted into the cloning site. The phagemid has the bacteriophage *f1* sites of initiation and termination. This allows for synthesis of single stranded DNA when an E. coli strain (XL1-B) is simultaneously infected with both the lambda vector and the f1 or M13 helper bacteriophage. The XL1-B bacteria that are infected by packaged phagemids can be plated on LB/ampicillin plates and will maintain the double stranded plasmid.

DNA sequencing

An adaptation of the dideoxy chain termination method (Sanger *et al.*, 1977) was used to determine DNA sequence. Sequence was obtained by using either double stranded DNA templates or single stranded DNA templates cloned into the M13 phagemid vector pBluescript SK(+). Sequencing reaction materials and protocols were obtained from a Sequenase® Version 2.0 DNA sequencing kit (United States Biochemical). Double and single stranded templates were prepared as described in Sambrook *et al.* (1989). The helper phage used to synthesize single stranded DNA was VCS-M13 (Stratagene). Sequencing reactions were loaded onto a 6% polyacrylamide gel prepared as described in Sambrook *et al.* (1989) and electrophoresed at 50 watts (constant power) for three to six hours in 1 X TBE buffer. Finally, the sequencing gel was fixed in two litres of 5% methanol, 5% acetic acid in water for 20 minutes. The gel was dried and autoradiographed as described in Sambrook *et al.* (1989).

BLAST search

The National Center for Biotechnology Information (NCBI) BLAST E-mail server allows users to send a specifically formatted e-mail message containing a protein sequence to the BLAST search server at the NCBI. A BLAST search is then performed against the specified database using the BLAST algorithm, and the results are returned in an e-mail message. In the case of the putative FEM-2 sequence, it was compared to all

the available proteins in the non-redundant protein database which includes SWISS-PROT, PIR(R), GenPept, and the GenPept updates (Altschul *et al.*, 1990).

Single worm PCR

The presence of plasmid extra-chromosomal arrays in transgenic lines was confirmed by using single worm PCR protocol as described by Williams *et al.* (1992) and Pilgrim and Bell (1993). Different primers were used for each transgenic line (see Primer list below). Amplification was performed using the following parameters: 95°C for five minutes, 61°C for one minute and 15 seconds, and 72°C for three minutes for one cycle, then 93°C for one minute and 15 seconds, 61°C for one minute and 15 seconds, and 72°C for two minutes 15 seconds for 29 cycles. The PCR machine that was used for all amplifications was a RoboCycler 40 (Stratagene). The reactions were then analyzed by electrophoresis on 0.7% agarose gels.

Heat shock of nematodes

In order to heat shock worms at specific times in development, single eggs were transferred to individual plates soon after they were laid. These eggs were then raised at either 20°C for N2 animals or 25°C for *fem-2(b245)*. The time that the egg hatched was recorded and the larva was allowed to develop for approximately two to three hours before it was shifted to 30°C (heat shock temperature). The worms were left at 30°C for various times, which were recorded, then shifted down to the original temperature. At the L3 - L4 stage these worms were examined for the dominant rolling phenotype that indicated they were carrying the desired heat shock construct, and those that were transgenic were left to grow (25°C for *b245* animals and 20°C for N2 animals) to adults to determine whether or not they were fertile.

Ethylmethanesulfonate (EMS) mutagenesis

Worms to be mutagenized were bleached to collect eggs (Sulston and Hodgkin, 1988) and then grown as a synchronous population to L4 stage. The worms were then collected in M9 buffer and treated as described in Sulston and Hodgkin (1988). The worms were mutagenized in a concentration of 0.1 M of EMS for four hours at 20°C. Following incubation in EMS, the worms were rinsed 4-5 times with M9 buffer and then dropped on seeded plates. The worms were allowed to recover for at least one hour then young healthy adults were picked to new plates and allowed to self-fertilize. The strain used and the selection technique is outlined in Figure 14. A typical screen involved setting up 120 P1 worms after mutagenesis on thirty, six cm plates (four worms/plate). These worms were allowed to self-fertilize at 25°C. Then 500 F1 progeny were transferred to fifty 10 cm plates (10 worms/plate). The progeny were collected in M9 buffer and then dauers were selected by exposure to a 2% SDS solution. Dauer larvae survive this treatment because they cease eating while non-dauers pump the SDS solution through their intestine and are digested from the inside out. The dauers were then allowed to recover at 15°C and examined for self-fertility. These plates were normally left at 15°C and checked six to eight days later for the presence of eggs.

Primer List

Name	Sequence 5' - 3'
fem-2 primers:	
(see Figure 4 for orientation and position of primers within the *fem-2* gene)	
AMC1	ACCTCGACGTCATAGCTAGT
AMC2	GACGACGATATGGTGAATAGG
AMC3	GATCACATCGGCATCGCCGTCG
AMC4	CATCAGTGACATCGGAATCTTC

DHA1/3/11	CAAAGATCTTGTCCCACCGAAGCCGGTAGTGG
DHA4(H1)	CGGTCTAGATGACGACGATATGGTGGATAG
DHA9(H2)	CTGCTAGAAGCTCCAGCGAGCTTGCGAAGC
TJ01	AACCACTACCGGCTTCGGTGGGACAGCTGG

Tc1 primer:

AMC5	AAAAGTGGATATCTTTTTGG

pBluescript primers:

-40	GTTTTCCCAGTCACGAC
REVERSE	AAACAGCTATGACCAT

All primers were synthesized in Ken Roy's laboratory, University of Alberta

Tc1 Insertion Screen

This screening procedure was taken from Rushforth *et al.* (1993) with the following modifications. Independent populations of MT3126 worms were established from 50 to 100 original adults on 100 x 15 mm plates consisting of approximately 8, 000-10, 000 animals. The worms were then collected in three ml of M9 buffer. Half of this was used for standard nematode DNA extraction and half used to propagate a fresh seeded plate that could be used to recover any insertion mutants that are isolated. DNA was isolated as described in Sulston and Hodgkin (1988). The remaining plates were wrapped in Para-film and incubated at 15°C. In addition, a fraction of this population was frozen at -80°C, as described by Sulston and Hodgkin (1988), for long term storage. In order to detect a Tc1 insertion in the *fem-2* gene, Polymerase chain reaction (PCR) was used on the genomic DNA isolated from these MT3126 populations. Primers that anneal outside the *fem-2* coding region (AMC2 and DHA3) were used in combination with a Tc1 primer (AMC5) located in the 54-bp terminal inverted repeats

(Rosenzweig *et al.*, 1983). The orientation and position of these primers is shown in Figure 4. Reaction parameters were used as described above in single worm PCR. The amplification products were electrophoresed and analyzed for the presence of a band corresponding to the total size of the region the primers flank (approximately 2.4 kbp).

Results

cDNA isolation and sequencing

An alignment of the physical and genetic maps for the left end of Linkage Group III suggested a location for the *fem-2* gene (Pilgrim, 1993). The *fem-2* gene was narrowed down to several overlapping cosmid clones by transformation rescue (Pilgrim *et al.*, in press) and subsequently to smaller plasmid subclones. Dr. David Pilgrim in this laboratory began to sequence plasmid subclones to determine the genomic sequence. In order to identify the 5' and 3' ends of the *fem-2* gene, as well as the position of the introns, a *C. elegans* cDNA library was screened for *fem-2* cDNA clones. This library was constructed in bacteriophage lambda ZAPII (Stratagene), and was generously provided by Robert Barstead (Barstead and Waterston, 1989). This library represents a population of cDNA clones from poly A^+ RNA isolated from a mixed stage N2 XX hermaphrodite culture. Approximately 86, 000 lambda clones were plated and plaque lifts onto nylon filters were performed. The filters were hybridized with the 15 kbp *Xba*I fragment from pDP#DBP003 (Figure 3). This fragment included more than the *fem-2* minimal rescuing region but was used in this screen because at the time, it was the shortest genomic clone that entirely contained the rescuing region. Subsequently, this minimal rescuing region was narrowed down to about 3 kbp (Pilgrim *et al.*, in press). In the first round of screening, a total of 43 cDNA clones were isolated that gave positive signals in hybridization reactions. By the *in vivo* excision protocol (Stratagene), phagemids (containing a functional bacteriophage *f1* origin of replication) were generated that consisted of the cDNAs cloned in the vector pBluescript SK(-) (Table 2). Preliminary sequencing of a subset of these clones showed that only four overlapping cDNA clones (pDP#AMc8, 19, 24 and 32) mapped entirely within the minimum rescuing region, the rest mapped outside this region. The remaining positive plaques were screened again using a smaller fragment, a 700 bp *Xho*I - *Sst*I fragment from

pDP#AMc19 (Table 2). This smaller fragment was used to identify cDNA clones that contained a 5' end that was as long as pDP#AMc19 or longer. This second screen identified five more cDNA clones (pDP#AMc3, 5a, 12, 14 and 18) of the putative *fem-2* gene.

The nine cDNA clones were partially sequenced from both the 5' and 3' ends using primers in the polylinker of pBluescript (-40 and reverse) by sequencing double and single stranded DNA. Additional sequence was obtained using various internal primers designated from known genomic and cDNA sequence (See Materials and Methods). A number of deletion subclones were also made from the cDNAs in order to obtain sequence other than the 5' and 3' ends (pDP#AMc12Δ*Bam*HI, pDP#AMc19Δ*Sst*I) (Table 2). The sequence data revealed that all the cDNA clones contained the same 3' end with a poly A^+ tail, but had different 5' ends and were possibly incomplete cDNAs. From the combined genomic and cDNA sequence, 1.8 kbp of contiguous cDNA sequence was constructed (see Figure 5). The complete sequence with the conceptual translation protein is shown in Figure 4 and will be used for subsequent references to specific nucleotides. Analysis of the largest open reading frame (ORF) demonstrated that two of the largest cDNAs (pDP#AMc8 and pDP#AMc12) contained 400 bp of predicted untranslated region and approximately 1350 bp of coding region. Two observations suggested that they were not full length cDNA clones; the ORF extended almost to the very 5' end of the cDNA sequence and a 22 base pair trans-spliced leader sequence (SL1 or SL2) that is trans-spliced on to the 5' end of 56% of all *C. elegans* transcripts (Spieth *et al.*, 1993) was not identified.

In many regions the sequence was very convincing because it had been determined six to eight independent times, from both DNA strands of genomic and cDNA clones. This provided a clear consensus and was assumed to be the correct

sequence. Other regions were ambiguous and multiple sequencing revealed no clear consensus. These discrepancies were resolved using a number of approaches. One experiment showed "CGCG" 1247 nucleotides downstream of the 5' end of the cDNA and two others showed "GGC" (missing the first C and a compression of the last CG). This was sequenced from both strands in opposite directions using double and single strand DNA as the template and was confirmed to be "CGCG". Another problem area was a specific eight base pair region (G-C rich) 105 nucleotides downstream of the 5' end of the cDNA (Figure 4) that exhibited a strong compression artifact. This was resolved in two ways; first, a new primer was synthesized (AMC1) which hybridized near the ambiguity, allowing wider spacing of bands closer to the bottom of the acrylamide sequencing gel, second, dITP was used in the sequencing reaction rather than dGTP. It has been observed that the substitution of dI for dG in DNA eliminates the secondary structures that produce gel artifacts (Gough and Murray, 1983). The use of dITP was able to resolve this sequence.

The final obstacle to the completion of the sequence of the *fem-2* gene was the identification of the 5' end of the transcript. This was determined by Dave Hansen in this laboratory using PCR based RACE (Rapid Amplification of cDNA Ends) (Frohman *et al.*, 1988). It was found that the 5' end had in fact been present on three of the cDNAs, with only a short segment of the SL1 leader sequence (see Figures 6 and 18). SL1 is a 22 nucleotide sequence that is trans-spliced to the 5' end of many mRNAs. This sequence is not present in the genomic DNA but seems to come from the 5' end of a 100-nucleotide RNA encoded by sequences in the spacer region between the 5S RNA genes (Krause and Hirsh, 1987). pDP#AMc8 and pDP#AMc12, the longest cDNA clones, contained three and eleven base pairs of the SL1 sequence respectively, that did not match the genomic sequence (Figure 18). In addition, a cDNA was isolated, pDP#AMc3, that shared the same 5' end as the other cDNAs and extended eight base

pairs into SL1 but was chimaeric and had a unidentified cDNA clone ligated to its 5' end (Figure 18). Sequence analysis of the cloned RACE product confirmed that the SL1 leader sequence was trans-spliced onto the *fem-2* mRNA. Further evidence that the 5' end of these cDNAs was in fact the 5' end of the transcript was that three of the other cDNA clones had 5' ends that were within 100 base pairs of the SL1 sequence. The complete cDNA sequence was searched for consensus sequences for donor and acceptor splice sites (Emmons, 1988) and four introns were found (Figure 6). The genomic sequence was completed and ambiguous sequences were confirmed after multiple sequencing attempts. The genomic and cDNA sequences were consistent with a single transcript of approximately 1.8 kb (Figures 4 and 6) that is seen on Northern Blots of RNA (Pilgrim *et al.*, in press). It was assumed that the first in-frame AUG codon corresponded to the initiation methionine, thus the open reading frame suggested a protein of 449 amino acids. There is a second in-frame AUG codon downstream of the first (in the second exon) which would reduce the number of amino acids in the protein by 34. A sequence that has been identified as a preliminary consensus for translation initiation in *C. elegans* is located just before the AUG initiation codon (Schneider *et al.*, 1986). The first AUG found in the *fem-2* cDNA matches 9/10 nucleotides from this potential "ribosome binding site" consensus. The second AUG sequence matches 8/10 nucleotides, therefore it is more likely that the first initiation methionine is utilized. In the 3' untranslated region (54 base pairs) two sequences that are typical of a *C. elegans* poly A signaling site were found. A sequence close to the consensus sequence, AAUAAA, (Emmons, 1988) was present 268 and 277 nucleotides downstream of the termination codon (Figure 4). The length of the poly A tail on each of the cDNAs was found to be different. The number of adenines added to the 3' end of the transcript ranged from 10 to 35. Also, one of the cDNA clones had a 3' end that was shorter than the other cDNA clones by three base pairs.

Analysis of the DNA sequence

The first BLAST homology search of the preliminary *fem-2* protein sequence was done in 1993 against all proteins in the database (Altschul *et al.*, 1990). The sequence identities that were found were not extensive but short regions appeared to be very similar. The protein that showed the highest sequence similarity to FEM-2 was the protein phosphatase 2C from rats. It was 36% identical and 56% similar over 155 amino acids. FEM-2 also showed amino acid similarity to a protein phosphatase from humans and rabbits, a yeast adenylate cyclase (CYR) as well as an elongation factor from a group of lower eukaryotes. The CYR protein exhibited 30% identity and 61% similarity over 72 amino acids and the elongation factor, 30% identity and 57% similarity over 46 amino acids. The identity with the adenylate cyclase and the elongation factor is outside the known active domains of these two proteins (Figure 7) (Kataoka *et al.*, 1985; Meester *et al.*, 1991). Because the sequence identity with these two proteins was relatively weak and did not overlap any catalytic domain, the alignment with the protein phosphatases was considered the most relevant. Also, the fact that many phosphatases are involved in signal transduction pathways (reviewed in Pallen *et al.*, 1992), similar to the proposed molecular model for sex determination, suggested a number of possible roles for FEM-2 (see Discussion).

A later BLAST search was performed (April 1995) using the final FEM-2 protein sequence to see if updated submissions were more similar to FEM-2. No extensive sequence identity was found beyond that discussed above. However, six separate regions of the FEM-2 protein (conserved motifs I - IV in Figure 8) exhibited significant similarity to an uncharacterized human ORF, as well as several mammalian and *Saccharomyces cerevisiae* protein phosphatases (PP2C) (see above paragraph). The search also detected sequence identity over these regions with the ABI1 gene from *Arabidopsis thaliana*, which is involved in the regulation of the plant's response to

abscisic acid (Leung *et al.*, 1994), and a kinase-associated protein phosphatase (KAPP) also from *Arabidopsis* (Stone *et al.*, 1994). When comparing the amino acid sequence of FEM-2 to the other proteins, the human ORF is 37% identical and 56% similar over the six motifs (190 amino acids), rat PP2C is 37% identical and 49% similar, and ABI1 is 31% identical and 46% similar. FEM-2 also showed protein sequence similarity to another predicted *C. elegans* protein F42G9.1, which is the first ORF from the cosmid F42G9 sequenced by the *C. elegans* Genome project (Sulston *et al.*, 1992; Wilson *et al.*, 1994). The sequence identity between FEM-2 and F42G9.1 is restricted to the conserved motifs, suggesting that these two proteins may be distantly related.

The amino-terminal third of the predicted FEM-2 protein showed no significant similarity to any protein sequence in the database. Most of the PP2C proteins to date have short (30-40 amino acids) amino termini (Cohen, 1989), whereas FEM-2, ABI1, and the human ORF have much longer amino terminal regions that extend 135-170 amino acids beyond conserved motif 1 (Figure 9). The extended amino terminal domains of the three proteins show no significant sequence similarity to one another, although all are rich in charged amino acids and proline. It is interesting to note that the mutation in the *b245* allele causes a change at a glycine residue that is conserved in all proteins shown in Figure 8. This sequence similarity between the predicted FEM-2 protein and numerous protein phosphatases has suggested that FEM-2 may have phosphatase activity. The results from this homology search have led to the proposal of a number of potential roles for FEM-2 in the determination of sex that will be outlined in the discussion.

Minigene construction and rescue

To confirm that the cDNAs confer *fem-2* function, a minigene, lacking all introns except the first, was constructed. If the last three introns or some enhancer sequence were essential for the proper expression or regulation of *fem-2* then it would not be

present in this construct and would not rescue the *fem-2* mutant phenotype. This plasmid was made by ligating a 1.4 kbp *Sal*I DNA fragment from pDP#AMc12 into a pBluescript plasmid containing 2.5 kbp of upstream genomic sequence (*Xho*I - *Sal*I fragment) as well as exon 1, the first 56 base pair intron and part of exon 2 (Figure 10). Because the 5' end of the minimal rescuing region lies only 240 base pairs upstream of the SL1 acceptor site it was expected that this plasmid would contain all the necessary promoter and enhancer sequences and be properly regulated. The cDNA fragment contained part of exon 2 and the rest of the 3' end of the cDNA, including the poly A+ tail and a portion of the polylinker. The cDNA fragment was cloned into both the sense (pDP#AM001(+)) and antisense (pDP#AM001(-)) orientations relative to the upstream promoter (Figure 10), with the assumption that AM001(+) would be transcribed and translated properly and that AM001(-) would not. When the potential translation from the antisense minigene was analyzed, stop codons were found in all three reading frames and the largest translation product that could be produced was 112 amino acids in length. This would not provide any functional FEM-2 product and therefore, should not rescue the *fem-2* mutant phenotype. These two plasmids were injected into the gonad of *fem-2(b245)* hermaphrodites, and the progeny were raised at 20°C (permissive temperature). Any transmitting rolling hermaphrodites were shifted to 25°C (non-permissive temperature) as L3-L4 animals to determine if the plasmid rescued the feminization effect of the *b245* mutation. If these rolling animals gave self-progeny then they had been rescued by the expression from the plasmid. In the case of AM001(+), 261 worms were injected, 39 rolling progeny were produced, and one stable transmitting line was obtained. With AM001(-), 135 worms were injected, 49 rolling progeny were found, and four transmitting lines were isolated.

The AM001(+) plasmid was able to rescue the *fem-2(b245)* phenotype and this strain could be maintained at 25°C because rolling animals remained self-fertile. Non-

rolling animals were self-sterile. Since the transformation rescue assay for *fem-2* function produces multicopy arrays (Mello *et al.*, 1991), a basal level of transcription from each gene may prove sufficient to rescue. Therefore, it is not known whether this minigene contains all the sequences responsible for wild type *fem-2* expression or regulation. The AM001(-) plasmid was not able to rescue the *b245* phenotype, as was expected, because rolling animals were not self-fertile at 25°C. These animals were tested in two ways; first, young (L2-L3) rollers were shifted to 25°C and checked to see if they laid eggs, and second, L4 or adult rollers were shifted to 25°C and left there to see if there were any F2 progeny. These L4/ adult animals had already produced sperm and therefore would have been able to fertilize their own oocytes after being shifted. Their progeny in turn, would not have inherited any *fem-2* product and would have been self-sterile. In both transgenic strains, the presence of the plasmid was confirmed using single worm PCR. Using a primer located in the polylinker of pBluescript and a primer in the *fem-2* gene, a fragment of the minigene was amplified and analyzed by electrophoresis (Figure 12).

The AM001(+) minigene was also injected into *fem-2(e2105)* hermaphrodites to determine whether it could rescue the XX germ line phenotype of the strongest allele and to check whether the somatic feminization of XO animals at 25°C could be rescued. The heterozygous *e2105* strain, DP62, was injected and its progeny were examined for the dominant rolling phenotype. After injecting 247 worms, one transmitting line was established. Rolling hermaphrodites were allowed to produce self-progeny and those that gave no dumpy animals in their progeny were assumed to be homozygous for *e2105*. Since these hermaphrodites were self-fertile, the minigene was able to rescue the XX germ line feminization caused by the *e2105* allele. However, this rescue did not allow the hermaphrodites to produce a wild type brood size and did not rescue all animals. Sixty percent of all rolling animals were self-fertile. The average brood size of

these transgenic hermaphrodites was 57 [wild type brood size is about 320 (Wood, 1988)]. The minigene that was present in *e2105* females did not cause maternal rescue in their non-rolling progeny. A transgenic adult hermaphrodite was shifted to 25°C and allowed to self-fertilize, and the XO progeny were examined for rescue. If the minigene did not rescue the XO phenotype, there would not be any phenotypic male progeny produced as they would be transformed into females. If the minigene did rescue, the only male progeny produced would be those that carried the array and therefore show the *rol-6* phenotype. Apparently, normal males were observed that had wild type male tails with most animals producing only sperm (Figure 11). A number of older males had a small number of oocytes in their germ line suggesting that the expression of FEM-2 was insufficient to prolong spermatogenesis throughout the adulthood of the worm. In addition, some of the males were not exhibiting the dominant roller phenotype implying that they did not carry the transgene. This variety in phenotype can be attributed to somatic mosiacism where varying number of cells inherit the extra-chromosomal array (reviewed in Fire, 1986).

Heat shock-minigene construct

The expression of *fem-2* was induced from an extra-chromosomal array at specific times during development in an attempt to rescue the *fem-2* mutant phenotype. If rescue was achieved, the next step would be to narrow down the heat shock period to a short window when *fem-2* is required. This would provide information on when exactly during development of the worm *fem-2* is necessary for proper male somatic development and spermatogenesis. The *fem-2* cDNA was ligated into a plasmid carrying the *hsp16-2* heat shock promoter (Stringham *et al.*, 1992). As well, the expression of antisense *fem-2* was examined in order to see if it could cause feminization of wild type N2 worms. The entire cDNA (1.8 kbp *Kpn*I/*Sma*I fragment from pDP#AMc8) was cloned into the *hsp16-2* plasmid in the sense orientation (Table 2). The antisense

plasmid was constructed in the same manner but with the 1.8 kbp *Kpn*I/*Sma*I fragment from pDP#AMc18 (Table 2). The "sense" construct was transformed into *b245* worms with 156 worms injected and six transmitting lines obtained. The "antisense" construct was injected into wild type N2 worms. In this case, 44 worms were injected and one transmitting line was established. Similar to the above mentioned transgenic lines, single worm PCR was performed to confirm the presence of the plasmid (data not shown).

The transgenic strain carrying the "sense" plasmid, pDP#AM002, was exposed to heat shock at various times during development to determine whether *fem-2* expression from the plasmid was capable of rescuing the mutant phenotype. Since the temperature sensitive period for *fem-2* in the hermaphrodite germ line is from L1 to mid-L2 larval stages (Kimble *et al.*, 1984), this time period was used as a basis for when the worms were heat shocked. *C. elegans* exhibits a typical heat shock response when exposed to elevated temperatures (29 - 35°C) (Emmons, 1988). Transgenic animals were exposed to heat shock for a range of four hours to 24 hours starting with animals that were early L1 stage (Figure 13). These were *b245* eggs raised at 25°C, then shifted to 30°C and back down to 25°C after the specified time. In other cases, L1 worms were shifted for a short period of time every two hours over a span of 10 hours (Figure 13). In all cases, these animals were not rescued as judged by their inability to produce any self progeny. Some worms did not survive the heat shock if it extended more than 10 - 12 hours as embryonic cell death occurs after prolonged exposure to 30°C (Stringham *et al.*, 1992). There are a number of possible explanations for why this heat shock construct did not rescue the *fem-2* phenotype. One is that the copy number of the extra-chromosomal array was not high enough to express sufficient levels of *fem-2* to rescue the feminization effect. The plasmid was present as shown by PCR results, therefore it was being transmitted properly to progeny (data not shown). Another explanation is that the promoter was not strong enough to express *fem-2* at a high enough level or that the

hsp16-2 promoter does not function in the germ line where it is expected that *fem-2* is expressed. The promoter has been shown to be active in most tissues (Stringham *et al.*, 1992) although it may not be functional in the gonad or the nonfunctional *b245* FEM-2 protein interfered with the heat shock induced protein. Perry *et al.* (1993) has used the *hsp16-2* promoter to induce expression of *her-1* cDNAs. They observed strong masculinization of XX animals after a heat shock of two hours at 33°C or one hour at 37°C twice a day (Perry *et al.*, 1993). Andrew Spence (personal communication) has also examined heat shock inducable expression of *fem-3* expression constructs. Overexpression of *fem-3* caused masculinization of somatic tissues in XX animals. These experiments demonstrate that the *hsp16-2* promoter can be useful to induce the expression of other sex determining genes.

The strain transmitting the "antisense" array, pDP#AM003, was analyzed in similar fashion to pDP#AM002. This construct, if expressed, was expected to cause feminization of wild type worms as it would presumably produce a transcript that could bind to wild type *fem-2* mRNA and prevent its translation (or affect its stability) (A. Fire, personal communication). Transgenic worms were exposed to heat shock in the same fashion as described above except that these worms were raised at 20°C and shifted back down to 20°C after the heat shock. The various time periods for heat shock were also similar and are shown in Figure 13. Of the plasmid carrying worms that survived the heat shock 20 out of 111 XX animals were apparently feminized, as shown by their inability to produce self-progeny. These feminized worms were exposed to 30°C for 20 - 24 hours starting at about two hours post-hatch. This result suggests that the antisense *fem-2* does feminize wild type worms but not in all cases. A total of 20 N2 animals that did not carry an extra-chromosomal array were treated with an identical heat shock and no feminization effect was seen, therefore the feminization seen in the pDP#AM003 transmitting animals can be argued to be due to the presence of the array. Male

transgenic animals were also examined under the Nomarski microscope for any feminization (lack of sperm production) and these appeared to be normal males, exhibiting a normal male germ line and soma.

fem-2 suppressor screen

To date, the only suppressors of *fem-2* are *fem-3 (gf)* mutations that suppress the germ line phenotype of *fem-2* but only in a *fem-2(b245)* background, and some *tra-1* mutations suppress only the somatic feminization. In order to identify other genes that may interact with *fem-2*, a mutagenic screen was performed. This screen was designed to identify mutations that suppressed the strongest *fem-2* allele, *e2105*. An outline of how this screen was performed is shown in Figure 14. This screen takes advantage of a closely linked gene *daf-7* (dauer forming mutant) (Riddle, 1988). Dauer larvae represent an alternative developmental stage that is entered when there is overpopulation or a lack of food (Wood, 1988). These animals cease to eat and defecate, develop a specialized cuticle and behavior, and can exist in this state until food becomes available or population density decreases. Dauer larvae can be easily selected due to their resistance to chemicals such as SDS. They do not ingest the chemical and are protected by their specialized cuticle. Briefly, heterozygous *fem-2(e2105)* animals are mutagenized and allowed to produce self-progeny (Figure 14). From the F1 progeny, wild type animals are selected and transferred to new plates. These animals may have a new suppressor mutation but it will be in heterozygous form. The F2 progeny are raised at 25°C, then collected and exposed to a solution of 2% SDS in order to select the dauer larvae. The dauers are allowed to produce self-progeny and their progeny are checked for self-fertility. The only animals that will be self-fertile will be homozygous for *fem-2* as well as the new suppressor mutation. If a dominant suppressor is isolated, animals that are heterozygous for the suppressor mutation will also be self-fertile. The resistance to SDS provides a simple selection scheme for dauer larvae and allows large numbers of mutagenized

worms to be screened. In addition to a suppressor, this screen could potentially produce a *fem-2* revertant. Over the course of approximately six months, a total of 6100 F1 animals were set up which produced progeny that were then screened for self-fertile animals. A number of dauers that were set up to recover at 15°C produced self-fertile progeny. These animals were backcrossed to the parental strain to confirm their genotype and were shown not to contain the original *fem-2* mutation. One quarter of the progeny from this cross should have been self-sterile because they were homozygous for the *fem-2* mutation but heterozygous for the putative suppressor. An unusually large number of worms in the F3 generation were fertile, suggesting that these animals did not have a real suppressor mutation but that something else was causing these animals to be self-fertile. In addition, approximately one quarter of the progeny from these animals exhibited a dumpy phenotype because they escaped the lethality of the *let(s2483)* mutation. Normally, worms homozygous for the *sC1* cross-over suppressor should be dead. These animals, therefore, were presumed not to be carrying a new suppressor mutation. One possible explanation for the large number of these self-fertile animals is that a small number of eggs present on the plate could have survived the SDS treatment. These animals could have been of any genotype and therfore, may have been heterozygous *fem-2* animals. There are a number of other possible explanations which originate from problems with the original strain. Since the *sC1* cross-over suppressor marked with a dumpy allele as well as a lethal mutation that seemed to be not 100% penetrant, some dumpy animals survived and were not eliminated from the screen. Another potential problem was that there could have been a cross over event between *fem-2(e2105)* and *daf-7(e1372)* as they are approximately one map unit apart. This is not a very likely event as *sC1* acts as a suppressor of recombination. The homozygous *daf-7* F2 animals would have been self-fertile after recovery from dauer. The frequency with which these self-fertile animals were recovered suggested that there was something wrong with the strain rather than the actual recovery of many new *fem-2* suppressors.

It is also possible that a genetic suppressor does not exist for the *e2105* allele and that one will not be found in a mutagenic screen.

Tc1 insertion screen

As in many other organisms, spontaneous mutations in *C. elegans* are often caused by the insertion of transposable elements. The *C. elegans* genome may contain five different families of transposable elements (for review see Berg and Howe, 1989). Tc1 elements are present in the genomes of all known strains, but the copy number and frequency of transposition vary widely. The mutator strain MT3126 carries a mutation, *mut-2(r459)*, which activates Tc1 transposition and excision at particularly high levels (Collins *et al.*, 1987). The technique of isolating genes using transposon tagging has been widely used (Hirsh *et al.*, 1979; Files *et al.*, 1983) and the screen used in this study was based on work done by R. Plasterk (personal communication). Insertions in exons identify potential null alleles. In order to isolate a null mutant of *fem-2*, a PCR-based screen was set up with the expectation of a Tc1 element inserting into the coding region of the *fem-2* gene. The ends of the Tc1 element contain a 54 base pair inverted repeat. Therefore, an oligonucleotide was constructed that would prime DNA synthesis away from the transposon in both directions (Figure 15). The strategy involves using PCR with pairwise combinations of the Tc1 primer and *fem-2* specific primers on genomic DNA isolated from the mutator strain. If a Tc1 element has inserted close enough to one of the *fem-2* primers, a fragment of DNA will be amplified that represents the region between the insertion site and the *fem-2* internal primer. Figure 15 illustrates what the insertion mutant DNA would look like and where the primers are located. This type of screen is very useful because it is independent of the phenotype of homozygous *fem-2* animals, since insertions of transposable elements can be detected in the heterozygous form.

Individual populations of the mutator strain, MT3126, were grown on large worm plates. The DNA was then isolated from a portion of the worms while the remainder were used to seed another plate or frozen in liquid nitrogen for long term storage. These are required in order to recover any insertion mutants that are obtained from the screen. This DNA was used in PCR reactions using *fem-2* primers located near the 5' and 3' ends of the coding region. Fifty-five separate populations were screened using this PCR technique and no Tc1 insertion mutant was found. Each population consisted of approximately 8, 000 - 10, 000 worms. A laboratory in the Netherlands that screens libraries of frozen populations as a service to the worm community has also attempted to isolate a *fem-2* Tc1 insertion mutant. Using primers from this laboratory, they were unsuccessful in obtaining an insertion mutant for *fem-2*. It is possible that the *fem-2* coding region does not contain the proper target sequence for the insertion event (Rozenzweig *et al.*, 1983). The other possible explanation for why an insertion mutant was not obtained was that an insufficient number of populations was screened. Rushforth *et al.* (1993) have obtained a Tc1 insertion in *mlc-2*, one of the two myosin light chain genes. This group screened a total of 102 independent populations of MT3126 and isolated four homozygous mutants. The frozen library that has been established may be useful in the future for any other genes cloned in this laboratory.

Discussion

Isolation and characterization of cDNA clones

After the location of *fem-2* was narrowed down to the left end of Linkage Group III to a 3 kbp fragment it was necessary to identify the *fem-2* transcript. The first step in further characterizing the *fem-2* gene involved screening a cDNA library. A number of cDNA clones were isolated from a cDNA library constructed by Bob Barstead (Barstead and Waterston, 1989) using a large genomic DNA fragment (15 kbp) as a probe. Forty-three positive clones were identified (from 86, 000 phage plated) in the first screen. Of these, 16 were partially characterized, by sequencing or a second round of hybridization selection. The remaining 27 were not characterized because they were not required to complete the sequence of *fem-2*. Of the 16 clones that were characterized, nine mapped to the *fem-2* minimum rescuing region. These *fem-2* cDNA clones were sequenced in order to determine the complete sequence of the *fem-2* gene. The seven cDNA clones that did not map to the *fem-2* region could have been cDNAs from other genes that are closely linked to *fem-2* including possible other members of a *C. elegans* operon (Spieth *et al.,* 1993). That is, similarly oriented genes are cotranscribed and individual mRNAs are formed by cleavage at the polyadenylation site and trans-splicing (to SL1 or SL2). An example of cotranscription is described by Spieth *et al.* (1993). The isolation of cDNA clones containing exons for two separate genes, *mai-1* (mitochondrial ATPase inhibitor) and *gpd-2* (glyceraldehyde-3-phosphate dehydrogenase) suggested that these genes are cotranscribed (Speith *et al.,* 1993).

Since nine out of the 16 cDNAs clones that were characterized mapped to *fem-2*, it could be extrapolated that 24 of the 43 clones isolated from the original screen are *fem-2* cDNAs. Twenty-four *fem-2* cDNAs isolated from a screen of 86, 000 phage suggests that the *fem-2* mRNA is a relatively abundant message. By comparing the

frequency of *fem-2* cDNAs to that of other sex determining genes, it suggests that *fem-2* is an abundant message relative to other genes. The cloning of *fem-1* involved screening 1.5 x 10^6 clones from a lambda gt10 library (S. Kim) which yielded over 100 positive cDNA clones (Spence *et al.*, 1990). In the case of *tra-1*, 2 x 10^6 phage from the lambda gt10 library were screened and about 120 positive clones were obtained (Zarkower and Hodgkin, 1992). This group was unable to identify any positive clones from more than 10^6 phage in the Barstead library. The isolation of cDNA clones for these two genes involved the screening of approximately 20 times as many phage compared to that screened for *fem-2*. If this many phage had been screened, about 480 *fem-2* cDNA clones would have been isolated. This suggests that *fem-2* is a relatively abundant message as it is approximately ten times more abundant than *fem-1* or *tra-1* and exhibits an absolute abundance of 0.03%.

None of the cDNAs was full length and this might be explained by the method in which the cDNA library was constructed. Briefly, RNA was isolated and the poly A^+ mRNAs were selected for using an oligo (dT)- cellulose column (Barstead and Waterston, 1989). Then, using reverse transcriptase and oligo (dT) primers, the complementary DNA strand is synthesized. The disadvantage of using an oligo (dT)-primed cDNA is that the reverse transcriptase must begin at the 3' end of the mRNA and may not reach the 5' end. This could be a possible explanation for the shorter cDNA clones that were obtained. Three of the cDNA clones extend to the 5' end of the mRNA but do not contain the full 22 nucleotide SL1 sequence. This may be a result of the stability of the trans-spliced leader sequence. Another possible explanation is that in the construction of the first strand, the reverse transcriptase will "turn the corner" and start to copy the DNA strand forming a hair pin loop. This loop can be digested by S1 nuclease treatment, producing a cDNA clone that is missing information from the 5' end of the mRNA (Sambrook *et al.*, 1989). Alternatively, SL1 may function as a 5' cap on the

mRNA transcript and could fold back on itself. If the reverse transcriptase reaches this point along the RNA, it may not be capable of synthesizing the very 5' end of the SL1 sequence. Trans-splicing presents an unusual phenomenon which has been shown to occur with a number of sex determining genes including *tra-2, tra-3, fem-3*, and *tra-1*. SL1 has been shown to be trans-spliced to about 56% of mRNAs in *C. elegans* (Speith *et al.*, 1993), and genes whose mRNAs are trans-spliced do not appear to share a common function. It has been proposed that spliced leader sequences may act as good translational initiation sites that are located near the AUG codon (Spieth *et al.*, 1993).

As mentioned previously, the first in-frame AUG gives rise to a protein of 449 amino acids and is assumed to be the initiation codon that is used for translation of *fem-2* mRNA. The second in-frame AUG, if used, would produce a predicted protein of 415 amino acids. However, there is no reason why the first AUG sequence should not be recognized by the ribosomal complex. Further evidence that the first AUG is used is that the sequence surrounding this codon matches a potential "ribosome binding site" (N N A/C A A/C A/C **A T G** N) consensus sequence more precisely than that surrounding the second AUG (Schneider *et al.*, 1986).

After sequencing the nine cDNA clones it was observed that the length of the poly A^+ tail on each cDNA varied. The number of adenines ranged from 10 to 35. This could be a result of how the cDNA library was constructed. When the first strand is synthesized using reverse transcriptase, the poly (dT) primer is a mixture of three primers (Barstead and Waterston, 1989). This mixture is dT (>25) with a dA, dC, or dG at the 3' end which ensures binding to the poly A^+ tail of the mRNA at the junction of the tail and the mRNA sequence. This technique prevents transcription of long stretches of the poly A^+ tail and increases the chances of synthesizing full length transcripts. Another possible explanation is that the length of the poly A^+ tail depends on the stage of the

worm that the RNA was isolated from. An example of this developmentally regulated polyadenylation is seen with *fem-3* mRNA (Ahringer and Kimble, 1991). It has been shown that the poly A+ tail of the *fem-3* mRNA decreases in length during development. The lengths of *fem-3* poly A+ tails were found to be 30 - 150 nucleotides in embryos, 30 - 100 in larvae, and 30 - 50 in adults (Ahringer *et al.*, 1992). It is unknown whether this phenomenon occurs with *fem-2* mRNAs because variations of 30 - 50 nucleotides in poly A+ tail length have not been analyzed on developmental Northerns. Another reason for the various poly A+ tail lengths could be the fact that the 3' end of the *fem-2* cDNA contains two potential polyadenylation signals. One signal may cause a shorter poly A+ tail to be added and the other, a longer poly A+ tail. The *fem-2* cDNA does not have the standard AAUAAA consensus (Wickens, 1990), but instead has two variants of that sequences, UAUAAA and UAUAAU (Figure 4). These variations do occur in other genes, though relatively infrequently (Sheets *et al.*, 1990). One of the cDNA clones, c19, has a 3' end that is three base pairs shorter than the rest. This could be a result of the two polyadenylation signals mentioned above. It is possible that the first signal causes polyadenylation at one position and the other signal causes polyadenylation three base pairs down stream of this. The 3' end of the cDNA is likely to be authentic because it produces a message of the appropriate size (when compared to that on a Northern blot) and the genomic sequence does not contain any long stretches of adenines that might have caused mis-priming of the cDNA.

The *fem-2* sequence that was derived from the cDNA and genomic clones predicted an ORF of 1350 base pairs (449 amino acids) as well as a 5' UTR of 54 base pairs and a 3' UTR of 394 base pairs. If the 3' and 5' UTRs from *fem-2* are compared with those from other sex determining genes, they share some common characteristics. First, most of these genes contain short 5' UTRs (35-73 base pairs) (Kuwabara *et al.*, 1992; Spence *et al.*, 1990; T. Barnes, personal communication), with *tra-1* and *fem-3*

having longer 5' UTRs, 150 and 250 base pairs respectively (Zarkower and Hodgkin, 1992; Ahringer *et al.*, 1992). Also, many of these genes have long 3' UTRs that have been implicated in regulation of the gene. A well studied example of this is *fem-3*, which contains a 5' UTR of 258 base pairs and a 3' UTR of 269. An inverted repeat sequence has been found in the 5' UTR of *fem-3* that may form a stem-loop structure and may inhibit translation (Ahringer *et al.*, 1992). The *fem-3* 3' UTR has also been shown to be necessary for the proper regulation of *fem-3* activity (Ahringer and Kimble, 1991). The 3' UTR sequence contains a potential binding site for one of the *mog* proteins (J. Kimble, personal communication). A regulatory element has been found in the 3' UTR of *tra-2* that is essential for negatively regulating *tra-2* at the translational level (Goodwin *et al.*, 1993). Therefore, it seems possible that the extended 3' UTR of *fem-2* may have some role in its function or regulation which is currently unknown.

fem-2 contains four introns, two that are relatively short, 50 - 60 base pairs, and two that are quite long, 300-375 base pairs. The two shorter introns are typical of most *C. elegans* introns (75% are between 45 and 59 base pairs in length, Emmons, 1988). All four introns are very rich in adenine and thymine, which is another common characteristic of nematode introns. The donor and acceptor splice sites matched the consensus sequence (Emmons, 1988) in all cases, could be predicted from the genomic sequence, and were later confirmed by comparison of cDNA and genomic sequences. The second intron contains an inverted repeat (Figure 16) which occurs 13 times. These repetitive elements may be the result of duplication of target sequences upon insertion of a transposable element (Kleckner, 1981). The fact that the minigene cassette (without intron 2) is able to rescue the *fem-2* mutant phenotype indicates that this inverted repeat sequence is not necessary for proper function of the gene. However, it must be acknowledged that rescue due to an extra-chromosomal array may not reflect the normal regulation of the gene.

The cDNA clones that were obtained from this library are only representative of hermaphrodite RNA because the library was constructed from a mixed stage hermaphrodite population (which may contain 0.05% spontaneous males, Hodgkin, 1983). Therefore, any male-specific mRNAs would not be present in the cDNA library and a male-specific *fem-2* transcript would have been overlooked. From the sequence that was obtained, it was clear that there was only one *fem-2* transcript and no alternative splicing was necessary for full *fem-2* activity in hermaphrodites. Two results that were obtained later in this project support the assumption that there is only one *fem-2* transcript. When RNA from a mixed population of males and hermaphrodites is analyzed on a Northern blot, only one RNA species is observed, suggesting that there is no male specific transcript (Pilgrim *et al.*, in press). *lac Z* staining has shown *fem-2* expression in XO animals throughout the adult stage of the nematode (T. Johnson, 1995). Therefore, if there was a male specific transcript, it should have been identified on the Northern blot of male and hermaphrodite RNA because *fem-2* activity seems to be required for a longer period in the development of males. It is possible that a rare transcript may not be identified on a Northern blot or be represented in the cDNA library. The temperature sensitive period for *fem-2(b245)* in XO animals also extends through adulthood (Kimble *et al.*, 1984) suggesting that it is required for maintaining spermatogenesis in the male.

Additional evidence suggesting that there is only one *fem-2* transcript is that the minigene cassette rescues all phenotypes of homozygous *fem-2(e2105)* XO animals at 25°C. The minigene is comprised of the 3' 1450 nucleotides of pDP#AMc12 attached to the amino terminus of *fem-2* and the complete *fem-2* promoter. Normally, *e2105* XO animals are transformed into fertile females at 25°C. Transgenic animals carrying the minigene show a normal male soma as well as the production of only sperm in the germ

line. These results suggest that the single transcript identified is sufficient to rescue the feminization effect of *fem-2(e2105)* on XO animals at 25°C. Once again, it should be considered that this is a transgenic rescue effect and may not necessarily represent the normal situation (i.e. over-expression of this construct may compensate for a lack of the other transcripts).

The minigene rescue studies are different than the injection of genomic clones for narrowing down the rescuing region, because the minigene does not contain introns 2 - 4. If the introns contain any sequences necessary for proper function of *fem-2*, this construct will not rescue the *fem-2* defect. This minigene plasmid was able to partially rescue the *fem-2(b245)*phenotype at 25°C and therefore contained all the necessary sequences to allow for limited spermatogenesis. This plasmid also rescued the *fem-2(e2105)* mutant phenotypes in most XX and XO animals. When the cDNA was cloned in the opposite orientation with respect to the promoter it was unable to rescue the *b245* defect and worms remained self-sterile at 25°C. These results indicated that the sequences contained in the minigene are necessary and sufficient for rescue of the *fem-2* defect. The fact that the cDNA rescued both germ line and somatic defects in XX and XO animals demonstrates that the introns do not contain any important sequences required for tissue specific *fem-2* function.

Heat shock studies

The minigene cassette was also cloned into a heat shock expression vector which contained the *C. elegans hsp16-2* heat shock promoter both in the 'sense' and 'antisense' orientations. These two constructs were transformed into *fem-2(b245)* and N2 animals respectively. Worms carrying the 'sense' construct were subjected to varying periods of heat shock at different times during development in order to determine specifically when *fem-2* expression is required to rescue the mutant phenotype. No definite rescue was

obtained from these experiments. Therefore, it seems that induced *fem-2* expression under the control of the heat shock promoter could not rescue the *fem-2* defect. Other experiments that have been performed using the *hsp16-2* have used a heat shock ranging from 31°C for 30 minutes (C. Kenyon, personal communication) to 37°C for two hours twice a day (Perry *et al.*, 1993). Perry *et al.*, (1993) showed that the *her-1* full length cDNA under the control of the heat shock promoter causes masculinization of the germ line and soma of XX animals. The expression of the *her-1* large transcript is necessary for male development in all tissues (Perry *et al.*, 1993). Andrew Spence (personal communication) has analyzed transgenic worms carrying the *fem-3* gene fused to the heat shock promoter and observed that over-expression of FEM-3 caused masculinization of somatic tissues in XX animals. The most plausible explanation for the *fem-2* construct not functioning was that the heat shock promoter used was unable to promote transcription in the germ line or, if it did, at a very low efficiency. Expression of chimaeric transgenes in the germ line has been found to be problematic (A. Fire, personal communication). There is no known reason for this inactivity, but it is possible that the germ line has specific mechanisms which prevent expression of transgenes either by transcriptional repression or post-transcriptional mechanisms. It is interesting to note that the *fem-2* mutation can be rescued by genomic copies of the gene (containing the endogenous promoter), but expression constructs with the same *fem-2* sequence (without the *fem-2* promoter) cannot rescue the germ line phenotype (Pilgrim *et al.*, in press). Most likely, only germ line specific promoters (such as *fem-2*) are active in the germ line.

A portion of the transgenic N2 worms carrying the 'antisense' construct were apparently weakly feminized. This was the expected result if the 'antisense' product annealed to the wild type *fem-2* transcript and prevented translation, therefore blocking spermatogenesis. In this case, it appeared that the heat shock promoter was partially active and due to mosaicism, not all animals exhibit the feminized phenotype. If the

extra-chromosomal array is not inherited in all cells, a fraction of worms will not express the *fem-2* transcript in the proper cells. This could also explain why the minigene did not rescue all *fem-2(e2105)* worms as mentioned above. One future experiment would be to establish a line of worms carrying the "antisense" plasmid that has been integrated into the genome. Exposure of worms to gamma rays will cause multiple breaks in the DNA. Through recombination, the extra-chromosomal array will then insert into the genome of the worm (Fire, 1986). This integration would allow for stable transmission of the plasmid to all cells and would eliminate somatic mosaic expression patterns. It may be possible that the feminization seen is due to the large number of copies of the non-functional *fem-2* gene, not the anti-sense RNA. A future experiment that may be done to answer this question is to perform reverse transcriptase PCR (RTPCR) on transgenic animals to determine whether they are producing "antisense" *fem-2* RNA. Recent evidence suggests that any truncated *fem-2* construct that is unable to rescue the phenotype may cause a dominant feminization of *fem-2(b245)* XO and XX animals at the permissive temperature (D. Pilgrim, unpublished results). It may be that the plasmid interferes with the normal *fem-2* function that the *b245* allele provides at 20°C. This may explain why some of the animals are feminized, however the feminization was seen in such a small proportion of worms that it seems unlikely.

Sequence analysis

An alignment search of the protein sequence databases using the Basic Local Alignment Search Tool (BLAST) program revealed a number of sequence similarities to the predicted FEM-2 protein. The highest similarity was to a number of serine/threonine protein phosphatases (Type 2C) from mammals (PP2C) and yeast. Additional proteins that were similar to FEM-2 were an uncharacterized human ORF, obtained from a random cDNA sequencing project (Nomura *et al.*, 1993), a gene from *Arabidopsis thaliana* that is involved in the plant's response to abscisic acid (ABI1) (Leung *et al.*,

1994), as well as a kinase-associated protein phosphatase (KAPP) from *Arabidopsis* (Stone *et al.*, 1994). The sequence similarity between the predicted FEM-2 protein and protein phosphatases suggests that FEM-2 might have phosphatase activity. One way in which the phosphatase activity of some of these homologous proteins was assayed using a fusion protein expressed in bacteria. Protein phosphatase activity is demonstrated by purifying the fusion protein, and then examining the dephosphorylation of radio-labeled, phosphorylated casein as a substrate (Stone *et al.*, 1994). A similar assay will be performed using FEM-2, once the protein has been purified.

Although FEM-2 shows the most sequence similarity with PP2C enzymes it appears that it is not an exact homologue of these mammalian proteins and represents a distinct class of protein phosphatase 2C homologues. As shown in Figure 9, the proteins that show sequence similarity with FEM-2 can be grouped into three classes. Class A proteins, which includes FEM-2, ABI1, KAPP, and the human ORF, have an extended amino terminal domain. Class C proteins include the yeast and mammalian PP2C isoforms and the F42G9.1 predicted protein, which have longer carboxy termini than the other classes. These proteins also have a conserved acidic nine or ten amino acid peptide that is present between motifs IV and V. The other group of proteins, Class B, lack the extended carboxy termini common in Class C proteins, but contain a 'Asp-Trp' dipeptide sequence at the end of conserved motif I. The presence of the long amino terminal region in the Class A proteins, of which FEM-2 is a member, suggests some novel function for this domain. This region may play some role in regulation of the protein's expression or be a region of interaction with another protein.

By analyzing the role of the above mentioned proteins that show sequence identity with FEM-2, a possible biochemical function for FEM-2 can be identified. The mammalian protein phosphatase type 2C was one of the first proteins shown to have

sequence similarity with FEM-2 however, it is one of the least characterized protein phosphatases. PP2C is one of the four major classes of serine/threonine-specific protein phosphatases in mammalian cells and is characterized by its dependence on Mg^{+2} and its insensitivity to okadaic acid (for review see Cohen, 1989). PP2C has a broad substrate specificity and is capable of dephosphorylating a number of enzymes that regulate metabolism and cell cycle but its activity is low *in vitro* compared to PP1 and PP2 (Cohen, 1989). Of the four classes of serine/threonine protein phosphatases, PP2C is the least well characterized. A number of enzymes that are involved in protein phosphorylation, such as protein kinase C, cyclic AMP-dependent protein kinase, and mitogen-activated protein kinase (MAPK), are very similar between mammalian and yeast cells, and it may be expected that the role of PP2C might also be similar between these systems.

A study of the role of PTC1, the *Saccharomyces cerevisiae* PP2C homologue, in signal transduction has revealed a possible regulatory function for PP2C. The disruption of PTC1 on its own has no clear phenotype, however, when in combination with a mutation in a tyrosine phosphatase gene, it exhibited a synthetic growth defect (Maeda *et al.*, 1993). This implied that the two phosphatases together were involved in a common signal transduction pathway. Recently, Maeda *et al.* (1994) have identified interactions between the PP2C yeast homologue (PTC1) and the two kinases (MAPKK and MAPK) that are involved in the yeast signal transduction pathway for glycerol synthesis. It is interesting to note that yeast carrying a deletion of *ptc1* grow normally at 30°C but grow very poorly at 37°C (Shiozaki *et al.*, 1994). These results suggest that the phosphatase is essential for normal growth at high temperatures. This temperature sensitivity may be compared to that seen in *C. elegans* sex determination. In the case of *fem-2*, the null allele, *e2105* feminizes the male tail only at 25°C but not at 20°C. The phosphatases that are required for their respective signal transduction pathways may have partial activity at

lower temperatures, but when subjected to high temperatures, they exhibit a stronger phenotype. Three MAPK homologues have been isolated in *C. elegans*. *sur-1* (suppressor of *ras*) is involved in the regulation of vulval development (reviewed in Kayne, 1994; Selfors and Stern, 1994). Another gene, *mpk-1*, has been identified as a MAPK by sequence similarity (Lackner *et al.*, 1994) and is involved in *ras*-mediated induction of vulval cells. This gene was cloned using degenerate primers corresponding to amino acid sequences conserved in all known MAP kinases, and PCR amplification. The third gene, *mpk-2*, which was identified in the same manner as *mpk-1* but has not been implicated in the vulval induction signal transduction pathway (Lackner *et al.*, 1994). No mutant phenotype has yet been associated with this locus however, if it is involved in sex determination its expected phenotype would be opposite to that of the *fem-2* phenotype. The *sur-1* and *mpk-1* alleles that have been identified are probably not null alleles, since it appears likely that a null mutation in these genes is lethal (Kayne, 1994; Selfors and Stern, 1994). There has been no PP2C homologue identified for the regulation of the MAPK cascade involved in the regulation of vulval development.

The proteins with the greatest sequence similarity to FEM-2 were the ABI1 gene from *Arabidopsis* and the human open reading frame. The function of the protein encoded by the human ORF is unknown therefore, it will not be discussed any further. The role of ABI1 is has been shown to be involved in the plants response to abscisic acid. The plant hormone abscisic acid (ABA) regulates a number of diverse physiological processes including seed maturation and stress responses (Leung *et al.*, 1994; Miller *et al.*, 1994). The *abi-1* mutant is able to grow in the presence of inhibitory amounts of ABA (Leung *et al.*, 1994; Miller *et al.*, 1994). The ABI1 gene product predicts a novel type of signaling enzyme as its amino terminal region shows sequence similarity to the conserved structure of a Ca^{2+} binding site (Leung *et al.*, 1994; Meyer *et al.*, 1994). The ABI1 carboxy terminal portion shows extensive sequence similarity to

protein phosphatases (PP2C) from rat and yeast, as does FEM-2. The ABI1 protein has been shown to possess phosphatase activity *in vitro* (Leung *et al.*, 1994; Meyer *et al.*, 1994). It has been suggested that ABI1 may negatively regulate phosphorylation in response to ABA-induced Ca^{2+} changes (Leung *et al.*, 1994). The *abi-1* mutation converts a glycine to an aspartic acid (amino acid 180 of ABI1; Fig. 9) that is flanked by the conserved residues Asp-Gly-His and Gly in a highly conserved region among the PP2C homologues. The amino acid substitution is located in the carboxy-terminal domain of the ABI1 protein which is predicted to be responsible for phosphatase activity. The *abi1* mutation causes a dominant gain-of-function phenotype which could be due to constitutive phosphatase activity in the presence of ABA (Leung *et al.*, 1994; Meyer *et al.*, 1994). The mutation could lead to an alteration in substrate affinity or specificity which would also explain the dominant phenotype. Further biochemical analysis of the wild type and mutant forms of ABI1 are required to clarify the nature of the mutation. The presence of the *abi1* mutation suggests that this region of the protein is probably functionally important. The *abi1* mutation has recently been introduced into the *fem-2* cDNA using *in vitro* mutagenesis and transgenic animals exhibit a dominant feminization phenotype. A portion of the males contain oocytes in their gonad (D. Hansen, personal communication). Therefore, this particular amino acid substitution (Gly to Asp at amino acid 206 of FEM-2; Figure 8) does not cause a dominant gain-of-function phenotype as it does in *Arabidopsis*. A dominant gain-of-function allele of *fem-2* would be expected to cause masculinization of hermaphrodites and has not yet been identified.

Recently, another protein has been identified in *Arabidopsis thaliana* that shows sequence similarity to FEM-2. The phosphatase, KAPP (kinase-associated protein phosphatase), interacts with the phosphorylated form of a serine-threonine receptor-like kinase, RLK5 (Stone *et al.*, 1994). The RLK5 gene encodes a protein that is similar to growth factor receptor kinases. KAPP contains three domains; an amino-terminal signal

anchor which is involved in the cytoplasmic orientation of a membrane-anchored protein (High and Dobberstein, 1992), a kinase interaction (KI) domain, and a phosphatase catalytic domain (Stone *et al.*, 1994). The carboxyl-terminal domain of KAPP, which contains the phosphatase homology region, shows significant sequence similarity with FEM-2 and the other proteins shown in Figure 8. The amino acid sequence of this protein, when compared to FEM-2, is 33% identical and 60% similar over 156 amino acids. When KAPP is compared to the other FEM-2 homologues, it fits into the Class A proteins because it has a long amino-terminal domain which includes the KI domain and the anchor signal sequence (Figure 9). The identification and characterization of KAPP which has been shown to interact with the protein kinase catalytic domain of RLK5 in *Arabidopsis* may help to elucidate the signal transduction pathways that involve serine-threonine receptor kinases. The KI domain may represent a catalytic domain that interacts with other proteins carrying amino acids that are potentially phosphorylated. This 238 amino acid domain is located 5' to the phosphatase domain and ends at amino acid 337 (Figure 8) which correlates with the end of Motif 1. The remainder of the KI domain shows no sequence similarity to any of the FEM-2 homologue proteins. Therefore, it seems unlikely that FEM-2 contains a similar sequence that interacts with a receptor kinase.

If FEM-2 has phosphatase activity, it is unknown as to what its target(s) may be. One possible candidate is TRA-1 which is predicted to be a transcription factor and has been shown to have a potential phosphorylation site (deBono *et al.*, 1995). TRA-1 is downstream of FEM-2 in the sex determination pathway (Figure 1), therefore genetic evidence suggests it may be one of the FEM proteins directly represses TRA-1. Currently, there is no molecular evidence to support this speculation. A number of gain-of-function alleles of *tra-1* have been shown to be associated with amino acid substitutions in a small region of the TRA-1 protein, which has been called the gain-of-

function (GF) region (deBono *et al.*, 1995). This GF region contains a potential glycogen synthase kinase-3 (GSK-3) phosphorylation site. The consensus amino acid sequence of the site is Ser-Val-Pro-Pro-Ser (Wang and Roach, 1993) where the first serine residue is phosphorylated by GSK-3 if the second serine has been previously phosphorylated. It is possible that this serine residue in the GF region is the target of a serine/threonine kinase as well as a serine/threonine phosphatase. If FEM-2 were to act directly on TRA-1, it would have to act on another target protein in the germ line in order to promote spermatogenesis. FEM-2 may dephosphorylate TRA-1 directly, or inactivate TRA-1 indirectly through FEM-1 and/or FEM-3.

It has been shown that *tra-2* is post-transcriptionally controlled by a regulatory element present in its 3' UTR (Goodwin *et al.*, 1993). *tra-2(gf)* mutations have been mapped to a perfect direct repeat located in the 3' UTR. These repeats are called DREs, (direct repeat elements), and are proposed to be bound by a direct repeat factor, DRF, which inhibits translation of *tra-2* (Goodwin *et al.*, 1993). The translational regulation of *tra-2* is thought to release the *fem* genes to allow the onset of spermatogenesis in hermaphrodites. Then inhibition of *fem-3* permits the switch to oogenesis (Barton *et al.*, 1987). How *tra-2* represses the *fem* gene products is not understood and it is possible that TRA-2 may act on FEM-2 to inhibit the activity of the other *fem* genes. However, recent experiments have suggested that it is FEM-3 that directly interacts with TRA-2 (A. Spence, personal communication). The sequence of the *tra-2* locus predicts that TRA-2 is an integral membrane protein with a carboxy-terminal intracellular domain (Kuwabara *et al.*, 1992), and FEM-3 is predicted to be a intracellular protein (Ahringer *et al.*, 1992). A carboxy-terminal fragment of TRA-2 has been shown to interact directly with FEM-3 (A. Spence, personal communication). Direct binding of FEM-3 to the cytoplasmic region of TRA-2 could be essential for the proper regulation of the *fem* genes in hermaphrodites. These findings suggest that it is not TRA-2 that is dephosphorylated by

FEM-2 directly. Two other proteins, FEM-1 and FEM-3, in the sex determination pathway may be more likely candidates for targets of FEM-2 activity.

Serine and threonine residues that are potential phosphorylation sites have been identified from the amino acid sequence of FEM-1 and FEM-3 and could be targets for FEM-2 phosphatase activity (Spence *et al.*, 1990; Ahringer *et al.*, 1992). Potential sites for cAMP-dependent kinase activity contain the amino acid sequence Arg-X-Ser/Thr where X is any amino acid and sites for casein kinase activity contain a Ser-X-X-Asp/Glu amino acid sequence (Ahringer *et al.*, 1992). Four potential sites have been predicted from the *fem-3* sequence and seven have been identified in the *fem-1* sequence. FEM-1 and FEM-3 are plausible candidates for proteins which might interact with FEM-2 because they are similarly required for male development and act at the same level of hierarchy in the sex determination pathway. The *fem* gene products may act sequentially, one positively regulating the next, or in a multi-protein complex to carry out their dual function in the germ line and the soma. In the model presented in Figure 17, FEM-2 is shown to act on FEM-1, which is thought to be capable of mediating protein-protein interactions through its six ankryin repeats (Spence *et al.*, 1990). The dephosphorylation of FEM-1 may allow it to bind and inactivate TRA-1 causing male somatic development. If FEM-2 does not dephosphorylate FEM-1, it would be unable to bind and inactivate TRA-1, therefore allowing the development of a female soma. FEM-1 may play a different role in the germ line once it is activated. Another possibility is that if FEM-2 and FEM-1 interact to control somatic sex determination and there may be another FEM-2 target that directs germ line sexual fate.

A potential protein for this role is FEM-3, particularly in the germ line. This is supported by the fact that *fem-3 (gf)* mutations suppress the germ line defect of *fem-2(b245)*. These gain-of-function mutations are the result of deletion of a small repeated

sequence in the *fem-3* 3' untranslated region (UTR) (Ahringer and Kimble, 1991; Evans *et al.*, 1992). The 3' UTR is thought to be necessary for negatively regulating *fem-3* in XX animals to allow for oogenesis. It is possible that FEM-2 dephosphorylates FEM-3 in order to activate it and promote spermatogenesis. A *fem-3(gf)* mutation causes continuous spermatogenesis, as long as there is normal FEM-2 to activate it. Temperature sensitive alleles of *fem-2* (e.g. *b245*, assumed to have partial function) are only partially able to activate *fem-3(gf)*, resulting in an intermediate phenotype between only spermatogenesis or only oogenesis. A complete loss of function allele of *fem-2* such as*(e2105)* is unable to activate *fem-3*, regardless of whether it is insensitive to negative regulation. This results in complete feminization of XX and XO animals. How the extended activity of *fem-3* is able to compensate for the *b245* mutation and allow for spermatogenesis is unknown. It may be that increasing the levels of FEM-3 increases the odds that a partially functional *b245* protein will be able to dephosphorylate and therefore activate FEM-3.

Experiments are presently being carried out in this laboratory to isolate antibodies against FEM-2 as well as to determine the target protein that is dephosphorylated by FEM-2. These experiments will add to the understanding of how somatic and germ line sex are determined. One problem with all of the above mentioned models is that if the role of FEM-2 is to remove phosphate groups to promote the male mode of development, some other protein must first phosphorylate the target protein. To date, no kinase has been found among the genes in the sex determination pathway. If an associated kinase exists, it is possible that mutations in it do not lead to a sexual transformation phenotype. It is possible that the kinase is constitutive, and essential in other signal transduction pathways, or it provides a redundant function. A genetic approach could be used to identify proteins that interact with FEM-2. Mutagenic screens for gain-of-function alleles, dominant alleles, as well as suppressors of *fem-2* might isolate potential

interactions between FEM-2 and other gene products. In addition, molecular approaches could be used to identify potential interactions involving FEM-2, such as the two hybrid system in yeast or coimmunoprecipitation of proteins that may directly interact with FEM-2.

Genetic screens

The genetic screen for germ line suppressors of *fem-2* was performed in order to identify any proteins that might interact with FEM-2 in the sex determination pathway. This involved mutagenizing heterozygous *fem-2(e2105)* worms and examining F3 progeny for the presence of fertile animals (see Results section and Figure 14 for details). Among the existing genes in the pathway, several alleles of each gene have been identified, suggesting that the standard genetic screens for mutants affecting sex determination have most likely reached saturation. Therefore, it is unlikely that a new sex determining gene with a simple phenotype would be discovered. However, it is possible that a specific gain-of-function allele of another gene (e.g. *fem-1*) may be isolated that would suppress the *fem-2* defect. Due to cross-over events between *fem-2* and the phenotypic marker, *daf-7*, a genetic suppressor of *fem-2* was not isolated from this screen. To avoid similar problems with future mutagenic screens, a more tightly linked marker might be used, such as *unc-45* (for uncoordinated). Another problem that was encountered was with the cross-over suppressor that carried lethal and dumpy mutations. Worms that were homozygous for this mutation should have been dead, however a small number of dumpy animals were observed on some plates indicating that this lethal allele was not completely penetrant. A more efficient lethal marker would be beneficial to any subsequent screens.

Another approach would be to use a parental strain that was heterozygous for *fem-2* over another phenotypic marker such as an *unc* mutation. This would make it

necessary to pick heterozygous wild type worms from a population that included "unc" worms. It is possible that a suppressor of *fem-2* exists and that it simply was not identified by this screen. It was expected that a *fem-3(gf)* mutant might have been picked up by this screen because it has been shown that several *fem-3(gf)* alleles suppress the germ line feminization caused by *fem-2* mutations. The problem with this expectation is that the *fem-3* gain-of-function mutations isolated to date only suppress the *b245* allele, which is a temperature sensitive allele. The *fem-2* allele used in this screen was *e2105*, which is a putative null allele (Pilgrim *et al.*, in press) and is not suppressed by the *fem-3(gf)* mutations. Perhaps a positive control screen could be performed using *b245* which should identify *fem-3(gf)* mutations

The goal of the Tc1 insertion screen was to obtain a null mutant of *fem-2* by searching for a transposable element insertion into the *fem-2* gene. A null mutant is important for understanding the role of *fem-2*, as a complete lack of FEM-2 may suggest some function that other alleles may not. Recently, DNA encoding the *e2105* allele has been sequenced and has been shown to contain a GC to AT transition which results in a premature stop codon. The situation is complicated because an additional mutation has been found (D. Hansen, personal communication) and the details of this are yet to worked out. The protein product produced by this mutant is only 214 amino acids long and is assumed to have no FEM-2 function. The XO somatic phenotypes of *e2105* worms have been compared with *e2105* in trans to a deficiency for the region and support the conclusion that *e2105* is a null (Pilgrim *et al.*, in press). There is a possibility that despite the above information, *e2105* is not a null. The fact that *e2105* animals raised at 20°C are not completely feminized suggests that either *fem-2* is not essential for male somatic development at lower temperatures or that the 214 amino acids encoded by *e2105* retains some partial function at lower temperatures. While the latter is

not very likely, comparison with a known null allele (such as a transposon insertion allele) may provide the definitive proof.

A mutator strain, MT3126, that has a high level of Tc1 element mobility (insertion and excision) was used to try to induce an insertion into the *fem-2* coding region, thereby, disrupting its function. This insertion mutant should be fertile as a heterozygote and would produce fertile progeny. These worms could be identified, out of a population of 8, 000 to 10, 000, by performing PCR on DNA isolated from the entire population. Although the frequency of Tc1 insertion into different genes varies widely, it has been shown that insertions into most genes occur at frequencies greater than 10^{-7} in this mutator strain (Collins *et al.*, 1987). Primers used in this PCR are located in a terminal inverted repeat located in the Tc1 element and at the 3' and 5' ends of the *fem-2* coding region. If the PCR amplifies a band of the proper size, the population of worms that the DNA was from can be retrieved from frozen stocks that represent a library of putative insertion mutants. The populations of worms that were screened to date, have not revealed any Tc1 insertion mutants for *fem-2*. This may be because not enough populations were screened, or that there is no Tc1 insertion target sequence in the *fem-2* region. A palindromic sequence of twelve nucleotides has been identified that may represent a Tc1 target sequence (Rosenzweig *et al.*, 1983). The *fem-2* coding region has been searched for this sequence and one potential target was found (Figure 4). This sequence does not match the consensus exactly and may not be a functional insertion site. A positive control was performed to determine whether the procedure used was sensitive enough. Approximately 20 worms carrying a known Tc1 insertion element (in the *unc-119* gene, M. Maduro, personal communication) were detected in a background of about 10, 000 MT3126 worms. It is assumed that more than one worm carrying a potential Tc1 element would be present on the plate because it would have produced a number of self progeny that would also carry the insertion.

Another laboratory has attempted a similar screen for a *fem-2* Tc1 insertion mutant and has also been unsuccessful (Ronald Plasterk, personal communication). One possible experiment that could be performed in the future is to screen these populations using different *fem-2* primers. If the primers used in this study were not specific enough, no insertion mutants would be detected or a false positive could be obtained. Another problem that was encountered was the presence of two Tc1 elements in close proximity, such that they amplified a DNA fragment between the two elements and not a *fem-2* specific fragment. This frozen library of worms may prove to be useful in this laboratory at a later time as it can be used to clone genes by transposon tagging, as well as isolate null mutants.

Summary

The *fem-2* sequence was obtained from a combination of sequences from nine cDNA clones as well as a number of genomic clones. The cDNA sequence demonstrates that *fem-2* contains a SL1 trans-spliced leader sequence, a 5' UTR of 54 base pairs, a single ORF of 1350 base pairs, and a 3' UTR of 394 base pairs. The amino acid sequence predicts a protein of 449 amino acids. It has been demonstrated that this is the *fem-2* gene using a number of different approaches. The predicted transcript size of 1.8 kb is consistent with that seen on a Northern blot (Pilgrim *et al.*, in press). A minigene lacking all introns except the first, was tested for its ability to rescue the *fem-2* mutant phenotype. This construct was able to rescue the XX and XO feminization caused by the *fem-2* mutation. A construct that contained the cDNA in the wrong orientation to the promoter did not rescue the *fem-2* mutant phenotype. In addition, the alleles of *fem-2* that have been sequenced all contain base pair changes within the cDNA sequence (D. Hansen, personal communication). These experiments have shown that the cDNA that was sequenced was able to confer *fem-2* function.

The *fem-2* protein predicted from a conceptual translation product shows sequence similarity to a number of proteins that appear to have phosphatase activity. This sequence similarity suggests that FEM-2 may act as a phosphatase in the signal transduction pathway that controls sex determination in *C. elegans*. How FEM-2 interacts with other protein products in the pathway and its target for dephosphorylation is unknown, however, a number of possible roles for FEM-2 have been discussed. A genetic approach was used in an attempt to identify genes that may interact with *fem-2*, specifically, suppress the *fem-2* mutation. A suppressor mutation was not isolated from this screen and it is possible that a genetic suppressor of *fem-2* does not exist. Although, one may in fact exist and was simply not picked up by this screen. At the onset of this project it was important to isolate a null allele of *fem-2*, and a screen was set up to isolate

a mutant that contained a transposon insertion in the coding region of *fem-2*. Populations of a mutator strain were screened using PCR in order to identify any worms that had a transposable element in the *fem-2* region. This experiment would have produced a mutant that had a complete lack of *fem-2* activity and might have provided some insight into the role of *fem-2* in sex determination. Such a mutant was not isolated from this screen. It may be that the *fem-2* coding region does not contain the proper sequences for insertion of this particular transposable element. Sequence analysis as well as genetic tests have recently shown that the *fem-2* allele, *e2105*, is in fact a null allele, therefore it is no longer a high priority to obtain a *fem-2* insertion mutant.

Tables

Table 1. *C. elegans* strains used in this study.

Strain [a]	Genotype
CB3535	*fem-2(b245ts)*
CB3538	*fem-2(b245ts); him-8(e1489)*
DP22	*unc-45(r450ts)fem-2(e2105) sC1* [*dpy-1(s2171)*]
DP47	*him-8(e1489)*
DP51	*fem-2(e2105)unc-45(r450)/sC1* [*dpy-1(s2171)*];*him-8(e1489)*
DP57	*fem-2(e2105)daf-7(e1372ts)/sC1* [*dpy-1(s2171)*]; *him-8(e1489)*
DP62	*fem-2(e2105)/sC1* [*dpy-1(s2171)*]; *him-8(e1489)*
DP73	*fem-2(b245ts)*; edEx15[b] [pDP#AM001(-);pRF4] isolate a
DP74	*fem-2(b245ts)*; edEx16 [pDP#AM001(-);pRF4] isolate b
DP75	*fem-2(b245ts)*; edEx17 [pDP#AM002;pRF4] isolate a
DP76	*fem-2(b245ts)*; edEx18 [" " ; "] isolate b
DP77	*fem-2(b245ts)*; edEx19 [" " ; "] isolate c
DP78	*fem-2(b245ts)*; edEx20 [" " ; "] isolate d
DP79	*fem-2(b245ts)*; edEx21 [" " ; "] isolate e
DP80	*fem-2(b245ts)*; edEx22 [" " ; "] isolate f
DP81	*fem-2(b245ts); him-8(e1489)*; edEx18 [pDP#AM002b;pRF4]
DP84	N2; edEx24 [pDP#AM003;pRF4] isolate a
DP85	N2; edEx25 [pDP#AM003;pRF4] isolate b
DP87	*fem-2(e2105)daf-7(e1372ts) /sC1* [*dpy-1(s2171)let (s2483)*]
DP108	*him-8(e1489)*; edEx25 [pDP#AM003;pRF4] isolate b
DP109	*fem-2(e2105); him-8(e1489)*; edEx36 [pDP#AM001(+);pRF4]
MT3126	*mut-2(r459); dpy-19(n1347)*

References

[a] CB strains constructed at the MRC laboratory in Cambridge, U.K.

DP strains constructed in this laboratory.

MT strain constructed by Collins *et al.*, 1989.

[b] Ex represents the reference number for the extra-chromosomal array.

Table 2. Recombinant plasmids used in this study.

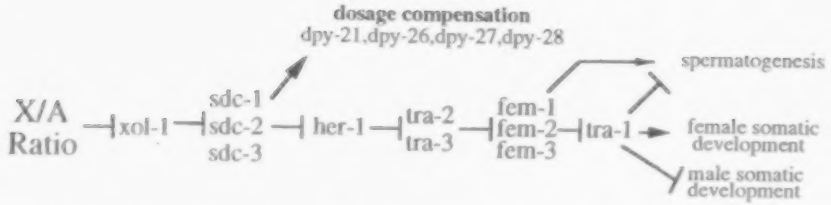
Plasmid	Cloned fragment
pRF4	4kpb *Eco*RI fragment of *C. elegans* genomic DNA containing the *rol-6(su1006dm)* collagen gene cloned into pBluescribe (Kramer *et al.*, 1990)
pDP#AMc3	chimaeric clone with 1.8 kbp *fem-2* cDNA fragment plus 1.5kbp of unknown origin in pBluescript SK(-)
pDP#AMc5a	1.3kbp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AMc8	1.8kbp *fem-2* cDNA fragment (full length) in pBluescript SK(-)
pDP#AMc12	1.8kbp *fem-2* cDNA fragment (full length) in pBluescript SK(-)
pDP#AMc12Δ*Bam*HI	1.1kbp derivative of pDP#AMc12 in pBluescript SK(-)
pDP#AMc14	1.7kbp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AMc18	1.75kbp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AMc19	1.3kbp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AMc19Δ*Sst*I	750bp derivative of pDP#AMc19 in pBluescript SK(-)
pDP#AMc24	550bp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AMc32	1.8kbp *fem-2* cDNA fragment in pBluescript SK(-)
pDP#AM001(+)	1.4kbp *Sal*I fragment of pDP#AMc12 cloned into pBluescript SK(-) ("sense" orientation) with 2.5kbp of upstream region
pDP#AM001(-)	1.4kbp *Sal*I fragment of pDP#AMc12 cloned into pBluescript SK(-) ("antisense" orientation) with 2.5kbp of upstream region
pDP#AM002	1.8kbp *Kpn*I/*Sma*I fragment of c8 cloned into the pPD 49.78 *hsp 16-2* expression vector ("sense" orientation)
pDP#AM003	1.8kbp *Kpn*I/*Sma*I fragment of c18 cloned into the pPD 49.78 *hsp 16-2* expression vector ("antisense" orientation)
pDP#DBP003	13kbp *Xba*I genomic clone in pBluescript SK(+)

pDP#DBP019	7kbp *Sst*I genomic clone in pBluescript SK(+) (does not rescue)
pDP#DBP023	9kbp *Xba*I/*Eco*RV genomic clone in pBluescript SK(+)
pDP#DBP026	6kbp derivative of pDP#DBP023 in pBluescript SK(+)
pDP#DBP141	10kbp *Xba*I genomic clone in pBluescript SK(+)

Figures

Figure 1. Genes involved in sex determination in *C. elegans*.

Model of genetic interactions between the genes involved in somatic and sex determination in *C. elegans*. Only the genes downstream of *xol-1* are illustrated, and genes affecting only the sex determination of the germ line (e.g. *fog* and *mog* genes) have been omitted. Pointed arrows signify positive regulation and barred arrows signify negative regulation. Adapted from Hodgkin, 1992.



XX soma	High	Low	High	Off	On	Off	On	female
XX germline	High	Low	High	Off	Off/On	On/Off	Off/On	sperm then oocytes
XO	Low	High	Low	On	Off	On	Off	sperm male soma

Figure 2. Molecular model for the determination of sex in *C. elegans*.

Model for the molecular interactions in the somatic sex determination pathway. Only genes that are important to the discussion of *fem-2* function are illustrated. See text for details. Adapted from Kuwabara and Kimble, 1992.

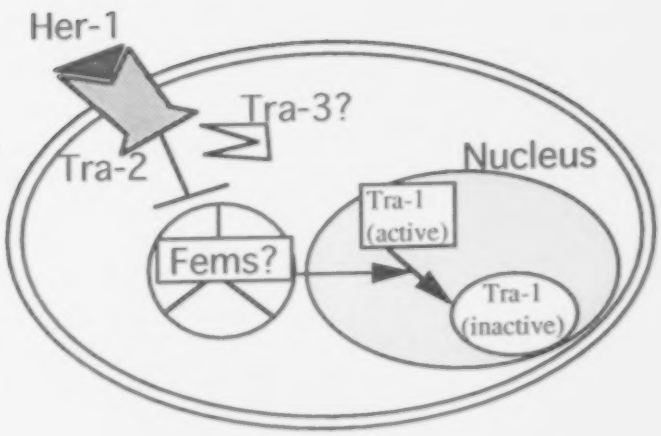

Her-1
Tra-3?
Tra-2
Nucleus
Tra-1
(active)
Fems?
Tra-1
(inactive)

Figure 3. Minimum rescuing region for *fem-2*

Restriction enzyme recognition sites are shown above the line. The most informative subclones that have been tested for their ability to rescue the *fem-2* mutant phenotype are shown on the left. "+" indicates rescue, "-" indicates no rescue. These results are data from Pilgrim *et al.* (in press) and illustrate the minimum region required to rescue the *fem-2* mutant phenotype.

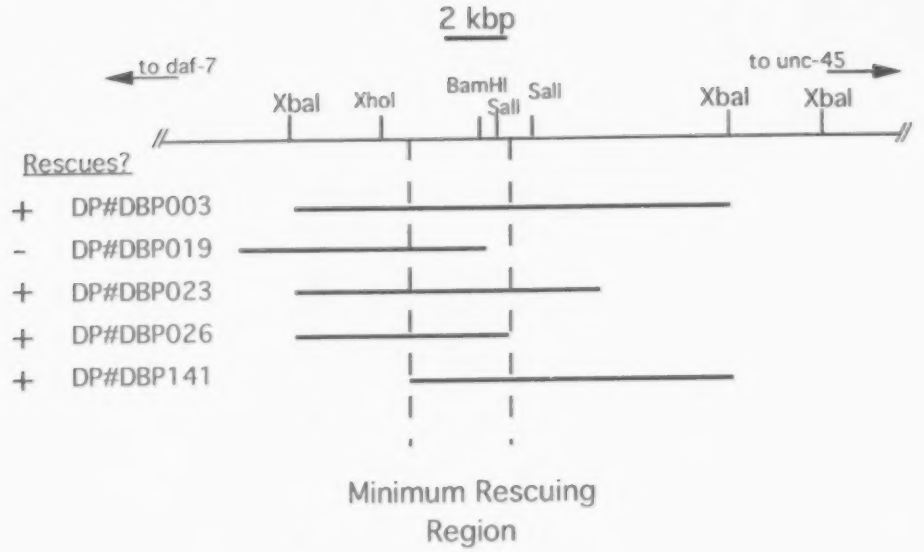

2 kbp
to daf-7
to unc-45
XbaI
XhoI
BamHI
SalI
SalI
XbaI
XbaI
Rescues?
+ DP#DBP003
- DP#DBP019
+ DP#DBP023
+ DP#DBP026
+ DP#DBP141
Minimum Rescuing Region

Figure 4. DNA sequence of *fem-2*.

Exons are shown in capital letters while introns and flanking regions are shown in lower case letters. The numbers refer to nucleotides of genomic DNA. Predicted translation of the sequence is shown above the sequence and the predicted initiation methionine is underlined. The primers are underlined and an arrow above or below the sequence indicates 5' to 3' direction of the primer. In addition, a number of relevant restriction recognition sites are underlined. Polyadenylation sequence is underlined in both strands of DNA at the 3' end of the sequence.

```
ttttttgatccatttttgatttgataaaaaatgttattaatacaatacaaatgcgaattaaattaatatttcgcgcgagaaacgcgcaacgcaccgaaagtgc   100
aaaaaactaggtaaaaaactaaacttatttttacaatattaatttatgtttatgtttacgcttaatttaattataaagcgcgctctttgcgcgttgcgtggctttccacg

                Hind III
gcaaacgagaagcttttgcgcgttttcggtgcagaaaatttctggcattatttttttttgttgaaatcccattgtttttttcctcgg                      200
cgtttgctcttcgaaacgcgcaaagaccacgtcttttaagacgtaataaaaaaaacaacaactttaggtaacaaaaaagagc

                                                                                                    M
ttttttctctgatttttatgttaaaaaacaatttattgaaatgtcagATCAACCCAGCTGTCCCACCGAAGCCGGTAGTGGTTTTAAAACTGGAAAAATACAT   300
aaaaagagactaaaaaatacaatttttgttaataacttttacagtcTAGTTGGGTCGACAGGGTGGCTTCGGCCATCACCAAAATTTTGACCTTTTTATGTA
                                                <<  TJ01

  E  K  V  N  E  E  R  D  A  V  F  E                    intron 1                                  (AMC3)
GGAAAAGTAAACGAGGAGCGGGATGCGGTTTTCGAGgtagaaaacgggatttttatatatttaatgcttagttagaaaaatgtatattccagGATCACATC                 D  H  I
CCTTTTTCATTTGCTCCTCGCCCTACGCCAAAAGCTCcatcttttgccctaaaatatataaatttacgatcaatctttttacatataaggtcCTAGTGTAG   400
     AMC3 >>                                                                 <<  AMC1
  G  D  R  R  R  S  V  R  S  L  L  E  E  A  F  A  D  E  M  E  K  T  S  Y  D  V  E  V  A  D  T  P  Q
GGCGATCGCCGTCGAAGTGTCCGGTCACTTCTGGAGGAGGCATTTGCAGATGAAATGGAGAAAACTAGCTATGACGTCGAGGTTGCCGACACTCCACAAC   500
CCGCTAGCGGCAGCTTCACAGGCCAGTGAAGACCTCCTCCGTAAACGTCTACTTTACCTCTTTTGATCGATACTGCAGCTCCAACGGCTGTGAGGTGTTG

  P  H  I  P  I  R  F  R  H  P  P  I  A  G  P  V  H  D  V  F  G  D  A  I  H  D  I  F  Q  K  M  M  K  R
CGCATATTCCAATACGTTTCCGTCATCCACCAATCGCCGGACCAGTTCATGATGTTTTTCGGAGACGCGATTTCACGACATTTTTCCAGAAAAATGATGAAAAG   600
GCGTATAAGGTTATGCAAAGGCAGTAGGTGGTTAGCGGCCTGGTCAAGTACTACACAAAAGCCTCTGCGCTAAGTGCTGTAAAAGGTCTTTTTACTACTTTTC
                Sal I
  G  Q  A  V  D  F  C  H  W  V  S  H  L  I  A  T  E  I  D  E  K  F  S  E  V  A  F  R  D  V  Q  Y  N
AGGCCAGGCGGTCGACTTTTGTCACTGGGTGTCTCACTTGATCGCCACAGAAATTGACGAGAAATTCAGTGAAGTTGCGTTCAGAGATGTTCAGTATAAAT   700
TCCGGTCCGCCAGCTGAAAACAGTGACCCACAGAGTGAACTAGCGGTGTCTTTAACTGCTCTTTAAGTCACTTCAACGCAAGTCTCTACAAGTCATATTA

  P  D  I  Y  V  T  D  S  T  T
CCTGATATTTATGTTACCGATAGTACCACAGgtaattattaaatgtgcaaaaatcttgcttattaatttagcagaccaagcatgggcctgctaaaatattccagc   800
GGACTATAAATACAATGGCTATCATGGTGTCcattaataaatttacacgttttagacgataattaaatcgtctggttcgtacccggacgatttataagtccg
```

```
                                    intron 2
aggcctatccccattggtcttgctaaatatttagcatacctaacttgggtcttgctaaatatttagcagaccatccccattggtcttgctaattatttagcag   900
tccggatagggggtaaccagacgatttataaatcgtctggattgaacccagacgatttataaatcgtctggtagggggtaaccagacgatttaataaatcgtc
                                    intron 2
accaacttatgggtcttgctaatttatttagcatacctaccatgagtcttgctaaatatttagcagacctat   1000
tggttgatacccagacgattaataaatcgtctggtggatacccagacgatttttataaatcgtatggatggtactcagacgatttataaatcgtctggata
                                intron 2
cctcattggtctgctaaatatttttagctggccattcccccattgatcttgctaatttttgcaaaaataaaaacttgttaaatagaatttttaaaaaaatcttc   1100
ggagtaaccagacgatttataaaaatcgaccggtaagggggtaactagacgatttaaaaacgttttattttgaacaatttatcttaaaaaatttttagaag
                                                E A K K L F N D K I W P A I D K I L Q Q
aaaaactcaacacttctgttcatcaatacaacttatttttcagAAGCCAAAAAGCTATTCAACGACAAAATTTGGCCGGCGATCGACAAGATTCTCCAGCA   1200
ttttgagttgtgaagacaagtagtttatgttgaataaaaagtcttCGGTTTTTTCGATAAGTTGCTGTTTTTAAACCGGCCGCTAGCTGTTCTAAGAGGTCGT
  N A E T C P I L S E K W S G I H V S G D Q L K G Q R H K Q E D R F
AAACGCCGAAACGTGCCCGATTCTGTCTGAAAAGTGGTCTGGAATCCACGTGTCGGGCGATCAACTGAAAGGCAACGTCACAAGCAAGAAGATCGGTTT   1300
TTTGCGGCTTTGCACGGGCTAAGACAGACTTTTCACCAGACCTTAGGTGCACAGCCCGCTAGTTGACTTTCCCGTTGCAGTGTTCGTTCTTCTAGCCAAA
                                                Bam HI
                                                D P I S
  L A Y P N G Q Y M D R G E                 intron 3
TTGGCGTATCCTAATGGGCAATATATGGATCGTGGAGAGgtttaggaggatttagGCctggaaaattactgaaaaatgaaatttttagGATCCAATTTC   1400
AACCGCATAGGATTACCCGTTATATACCTAGCACCTCTCCaaatccctcctaaaatccggaccttttaatgactttttacttttaaaaatccTAGGTTAAAG
                                        e2105=A transition
  V L A V F D G H G G H E C S Q Y A A G H L W E T W L E V R K S R D
GGTGTTAGCGGTGTTCGATGGTCACGGTGGTCACGAGTGCTCTCAGTACGCAGCCGGGCACCTTTGGGAGACATGGTTGGAGGTTCGAAAATCGCGACAT   1500
CCACAATCGCCACAAGCTACCAGTGCTCACGAGAGTCATGCGTCGGCCCGTGGAAACCCTCTGTACCAACCTCCAAGCTTTTAGCGCTCTA

  P S D S L E D Q L R K S L E L L D E R M T V R S V K E C W K G G S
CCTTCAGATAGCCTCGAAGATCAGCTTCGCAAGTCGCTGGAGCTTCTCGACGAACGAATGACTGTCAGAAGTGTGAAAGAGTGTTGGAAGGGTGGAAGTA   1600
GGAAGTCTATCGGAGCTTCTAGTCGAAGCGTTCAGCGACCTCGAAGAGCTGCTTGCTTACTGACATTCTTCACACTTTCTCACAACCTTCCCACCTTCAT
                                <<  DHA9
                                Hind III                                Eco RI
  Pvu II                                                              E F R Q
T A V C C A I D M D Q K L M A L A W L G D S P G Y V M S N I
CAGCTGTCTGCTGTGCAATCGACATGGATCAGAAGCTTATGGCGCTGGCGTGGCTCGGTGACTCACCCGGATACGTCATGTCGAACATTGAATTCCGTCA   1700
GTCGACAGACGACACGTTAGCTGTACCTAGTCTTCGAATACCGCGACCGCACCGAGCCACTGAGTGGGCCTATGCAGTACAGCTTGTAACTTAAGGCAGT
```

```
  L  T  R  G  H  S  P  S  D  E  R  E  A  R  R  V  E  E  A  G  G  Q  L  F  V  I  G  G  E  L  R  V  N
GTTGACACGTGGACACTCGCCATCCGATGAACGGGAAGCTCGGCGTGTTGAGGAAGCTGGCGGCCAGCTTTTTGTGATTGGAGGCGAATTGCGGGTGAAC  1800
CAACTGTGCACCTGTGAGCGGTAGGCTACTTGCCCTTCGAGCCGCACAACTCCTTCGACCGCCGGTCGAAAAACACTAACCTCCGCTTAACGCCCACTTG

                          b245=A transition    Xma III
 G  V  L  N  L  T  R  A  L  G  D  V  P  G  R  P  M  I  S  N  E  P  E  T  C  Q  V  P  I  E  S  S  D
GGAGTTCTAAATCTCACCCGAGCACTCGGAGATGTTCCTGGACGGCCGATGATCTCTAATGAACCGGAAACGTGTCAGGTGCCTATTGAAAGCTCGGATT  1900
CCTCAAGATTTAGAGTGGGCTCGTGAGCCTCTACAAGGACCTGCCGGCTACTAGAGATTACTTGGCCTTTGCACAGTCCACGGATAACTTTCGAGCCTAA

 Y  L  V  L  L  A  C  D  G  I  S  D  V  F  N  E  R  D  L  Y  Q  L  V  E  A  F  A  N  D  Y  P  V  E
ATCTGGTGCTCCTGGCTTGTGATGGAATTTCTGATGTTTTCAACGAGCGTGACCTGTACCAGTTGGTGGAGGCATTTGCGAATGATTATCCTGTTGAAGg  2000
TAGACCACGAGGACCGAACACTACCTTAAAGACTACAAAAGTTGCTCGCACTGGACATGGTCAACCACCTCCGTAAACGCTTACTAATAGGACAACTTCc
                                                potential Tc1 insertion site
                          intron 4
ttaataaaaattcttggagaaaaactaaatttccaaagccaaaaatcatttttcgttttttttttctgaaaaaaatcagaagattggctatttctcggg  2100
aattatttttaagaacctctttttgatttaaaggtttcggtttttagtaaaaaagcaaaaaaaaaagactttttttagtcttctaaccgataaagagccc

                          intron 4
ggttttcaaaaaaaaaaagtaaagattttcgaattattgtatttttttaaaaaagaaaatcgtgaatttactaaaatttgaaaaaatcaatttttaaaaaa  2200
ccaaaagtttttttttttcatttctaaaagcttaataacataaaaaattttttctttttagcacttaaatgattttaaacttttttagttaaaaatttttt

                          intron 4                                                                  D
accattcaaaattatttttaaggtaaaaaacattagccgacgttactttatttttttaatgtagaaaattctttaaaaatcgaaaaaaaatttttttcagAC  2300
tggtaagttttaataaaaattccatttttttgtaatcggctgcaatgaaataaaaaattacatcttttaagaaatttttagctttttttttaaaaaagtcTG

      Sst I                                                                                .
 Y  A  E  L  S  R  F  I  C  T  K  A  I  E  A  G  S  A  D  N  V  S  V  V  I  G  F  L  R  P  P  Q  D
TACGCGGAGCTCTCCCGTTTCATCTGTACAAAAGCGATTGAGGCTGGAAGTGCCGACAACGTGTCTGTTGTGATTGGCTTTTTGCGTCCACCGCAGGATG  2400
ATGCGCCTCGAGAGGGCAAAGTAGACATGTTTTCGCTAACTCCGACCTTCACGGCTGTTGCACAGACAACACTAACCGAAAAACGCAGGTGGCGTCCTAC

    Hind III
V  W  K  L  M  K  H  E  S  D  D  E  D  S  D  V  T  D  E  E
TGTGGAAGCTTATGAAACATGAATCAGACGATGAAGATTCCGATGTCACTGATGAGGAATAACTGCTTTTCGGTGGAAATTTTGCCTGAAAATTGGGAAA  2500
ACACCTTCGAATACTTTGTACTTAGTCTGCTACTTCTAAGGCTACAGTGACTACTCCTTATTGACGAAAAGCCACCTTTAAAACGGACTTTTAACCCTTT
                                  << AMC4

ATTCTTGCTAAACTTCGGTTGATTTTCTCAAATTTTTGTCGTTTTTTTTCCTAATTTTCGGTCGATTTTATCAAATTTTCCCCTGTTTTCCTGTGTCTAT  2600
TAAGAACGATTTGAAGCCAACTAAAAGAGTTTAAAAACAGCAAAAAAAAGGATTAAAAGCCAGCTAAAATAGTTTAAAAGGGGACAAAAGGACACAGATA
```

```
TTGTTTTTTGGGTATTCCAATTTTCTCTCTCCTATTCACCATATCGTCGTCATTTCTCCCCGTTATCTAGCAGCTTTCCGTGACCCACACACATATTTTT  2700
AACAAAAAACCCATAAGGTTAAAAGAGAGAGGATAAGTGGTATAGCAGCAGTAAAGAGGGGCAATAGATCGTCGAAAGGCACTGGGTGTGTGTATAAAAA
                                                << DHA4
                                    << AMC2
                                          Bst XI
GTTCTCGGTGTTCCTGTAGCATTTCACATTTACATACGTTTCCAACCAGTTGGCATCCCACCAAAAATCCGAGTACTTTCCTTCTCCTCCTCCACTTGAA  2800
CAAGAGCCACAAGGACATCGTAAAGTGTAAATGTATGCAAAGGTTGGTCAACCGTAGGGTGGTTTTTAGGCTCATGAAAGGAAGAGGAGGAGGTGAACTT

CATATATTTGATTTGAATTTATTTTATGATATAAATTGTATAATGACGTGGAAttactgggatattgagtagctgaataggttttgagacacttttgaca  2900
GTATATAAACTAAACTTAAATAAAATACTATATTTAACATATTACTGCACCTTaatgaccctataactcatcgacttatccaaaactctgtgaaaactgt

ctgatttcataatgactgtaactttcaactcgtgaaaatgtaaaaatcggagtataccagtcattttcagtagaaaatctccttttcaattgaaaatccg  3000
gactaaagtattactgacattgaaagttgagcacttttacatttttagcctcatatggtcagtaaaagtcatcttttagaggaaaagttaacttttaggc

ggatttttcgaaagaaaagctcgaaaacaccgcgaaaaaatgattaaaaaagtcatattttcaagagggaatttcaaatttcccgcttaca
cctaaaaagctttcttttcgagcttttgtggcgctttttactaatttttttcagtataaaagttctcccttaaagtttaaagggcgaatgt
```

Figure 5. Construction of cDNA sequence.

A partial restriction map of the rescuing region is shown with the genomic and cDNA sequences that were obtained below. This contiguous sequence from various cDNA clones was used to compare cDNA sequence to genomic sequence and to confirm that the entire gene was sequenced. Genomic clones are numbered and are represented by thin lines. cDNA clones are indicated by the "c" prefix and by thicker lines. Intronic regions are indicated by dashed vertical lines. Not all genomic sequences were included as they provided redundant sequence. The full identifiers for the genomic and cDNA clones is as follows: the genomic clones shown are classified as pDP#DBPXXX and were constructed by David Pilgrim and the cDNA clones are classified as pDP#AMXXX as shown in Table 2.

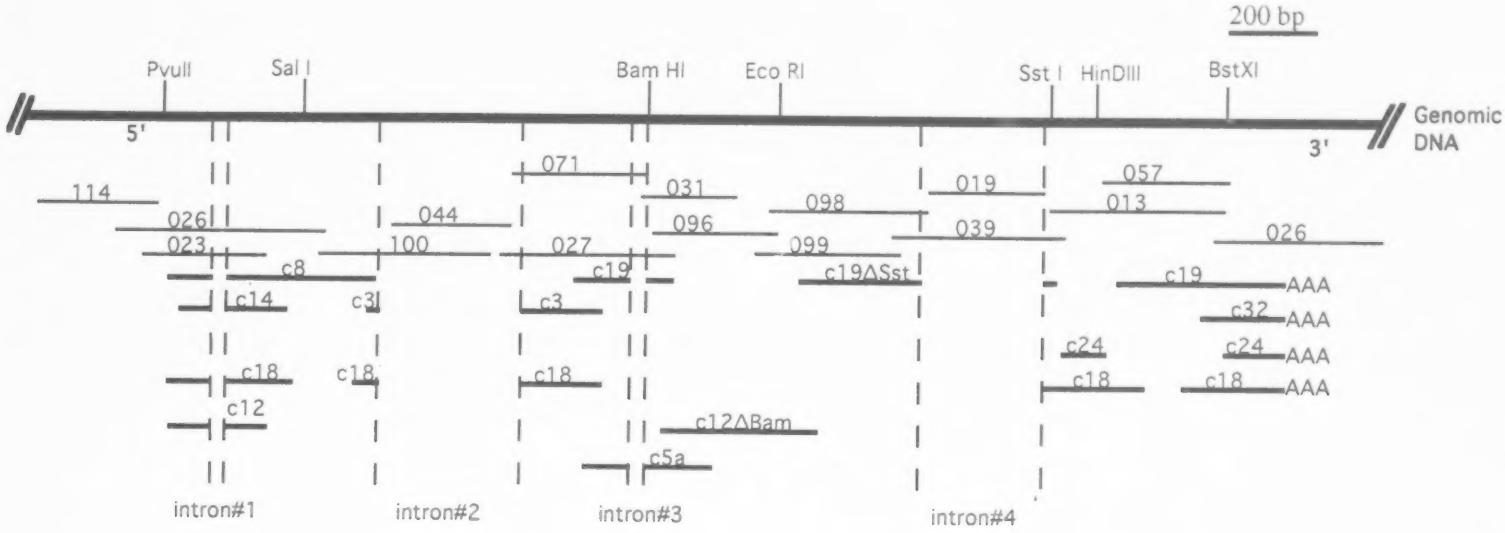

200 bp
PvuII
Sal I
Bam HI
Eco RI
Sst I
HinDIII
BstXI
5'
3'
Genomic DNA
114
026
023
c8
c14
c18
c12
071
044
100
027
c19
c3
c18
031
096
c12ΔBam
c5a
098
099
c19ΔSst
019
039
057
013
026
c19
AAA
c32
AAA
c24
c24
AAA
c18
c18
AAA
intron#1
intron#2
intron#3
intron#4

Figure 6. *fem-2* cDNA/genomic arrangement.

The sequence was searched for intron splice consensus sequences and arranged to show the exons and introns. The conceptually spliced product (1.8 kb) is shown with the SL1 and untranslated regions indicated with a shaded box. The largest ORF is indicated by the black region.

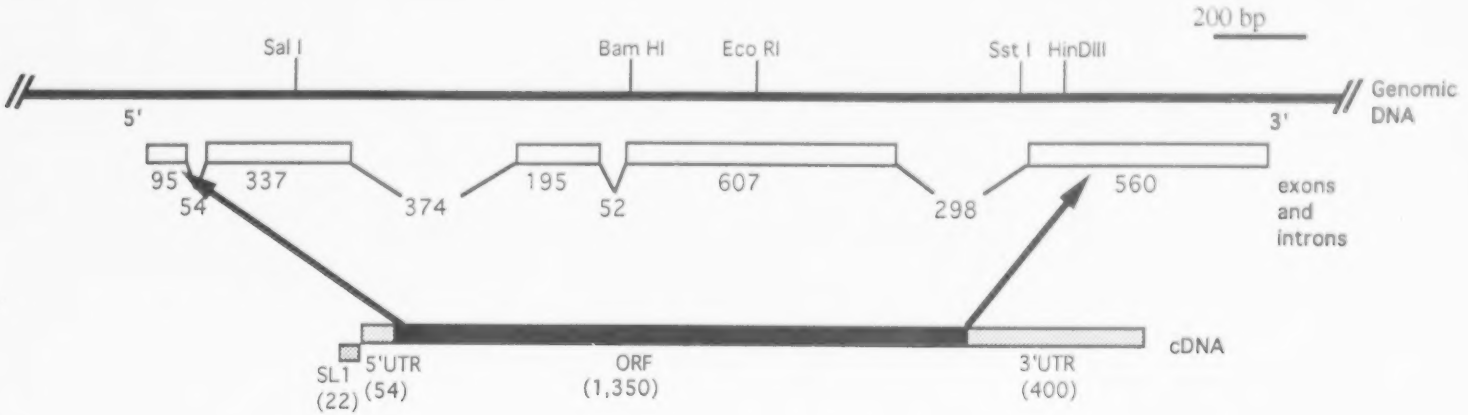

200 bp
Sal I
Bam HI
Eco RI
Sst I
HinDIII
Genomic DNA
5'
3'
95
54
337
374
195
52
607
298
560
exons and introns
cDNA
SL1 (22)
5'UTR (54)
ORF (1,350)
3'UTR (400)

Figure 7. *fem-2* homologue alignment.

The initial database search showed that a number of PP2C proteins, the yeast adenylate cyclase, and an elongation factor were similar to FEM-2. These proteins are shown with the region of similarity indicated by the lightly shaded boxes. The known catalytic domains of the proteins are also illustrated. The protein sequences and the relative catalytic regions were obtained from Kataoka *et al.*, 1985 (adenylate cyclase), Meester *et al.*, 1991 (elongation factor) and Wenk *et al.*, 1992 (PP2C from rats).

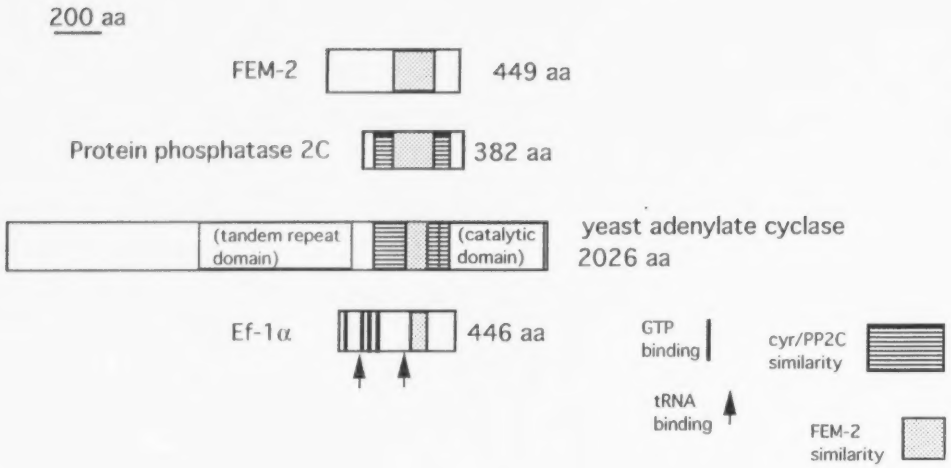

200 aa
FEM-2
449 aa
Protein phosphatase 2C
382 aa
(tandem repeat domain)
(catalytic domain)
yeast adenylate cyclase
2026 aa
Ef-1α
446 aa
GTP binding
cyr/PP2C similarity
tRNA binding
FEM-2 similarity

Figure 8. Sequence comparison of conserved motifs.

FEM-2 is aligned with its putative homologues showing the six conserved regions. Sequences were obtained from Leung *et al.*, 1994 (ABI1), Nomura *et al.*, 1993 (human ORF), Maeda *et al.*, 1993 (*Saccharomyces cerevisiae* PTC1), Wenk *et al.*, 1992 (PP2CB from rats), Sulston *et al.*, 1992; Wilson *et al.*, 1994 (CeF4, F42G9.1 protein from *C. elegans*), Stone *et al.*, 1994 (KAPP)." *" means >57% of the proteins have the same amino acid. " + " means >80% of the proteins have one of two amino acids. " · " represents a gap that has been introduced to increase the alignment. Numbers refer to the amino acid number at the beginning of the conserved motifs. The *b245* mutation is indicated by an arrow which changes a glycine to an arginine (Pilgrim *et al.*, in press). The *abi1* mutation is also indicated by an arrow which changes a conserved glycine to an aspartic acid (Meyer *et al.*, 1994; Leung *et al.*, 1994). Sequence accession numbers; FEM-2: U29515; ABI1: gpX77116; HORF: gpD13640; PTC1: spP35182; P2CB: spP35815; CeF4: gpU00051; KAPP: gpU09505.

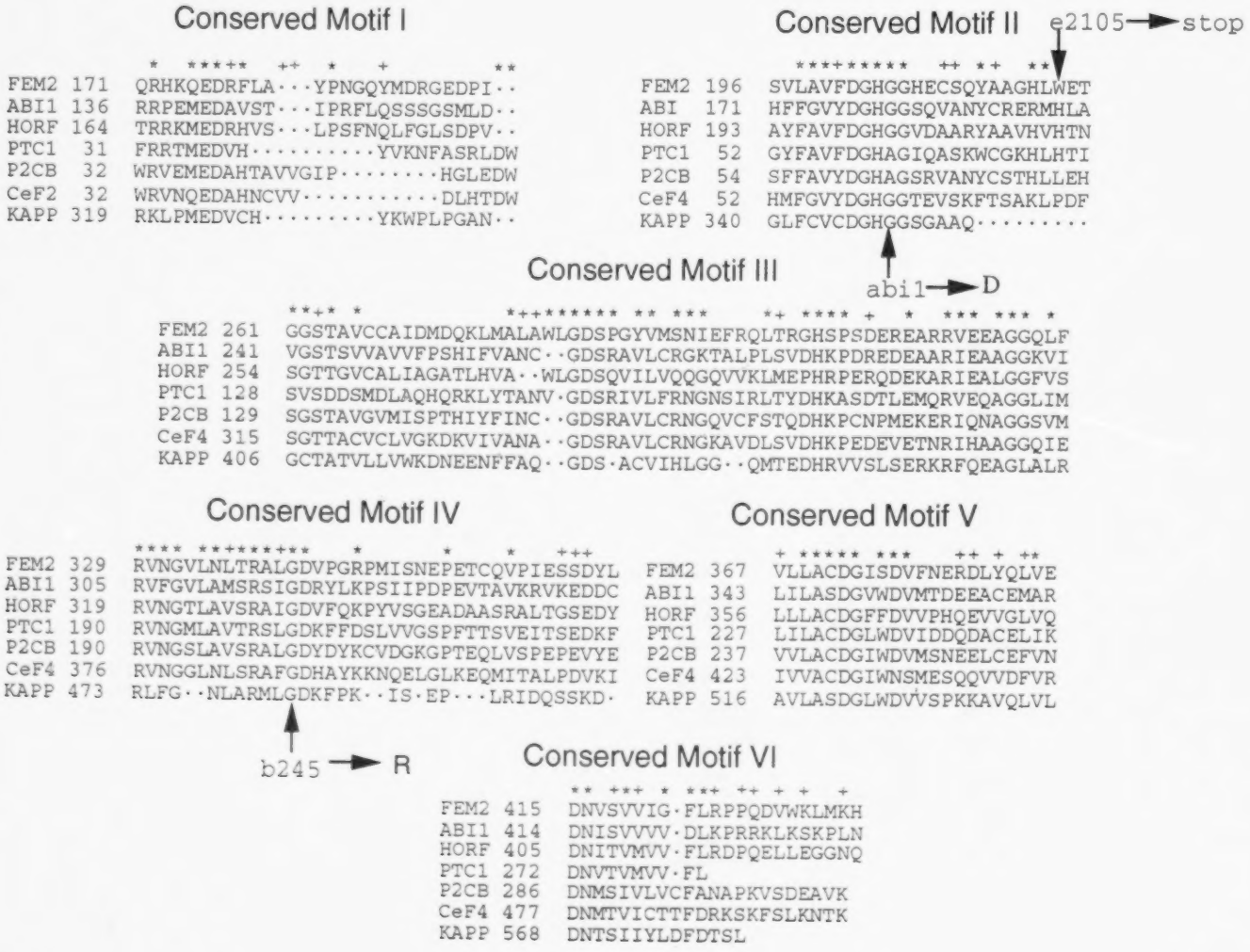

Conserved Motif I
FEM2 171 QRHKQEDRFLA···YPNGQYMDRGEDPI··
ABI1 136 RRPEMEDAVST···IPRFLQSSSGSMLD··
HORF 164 TRRKMEDRHVS···LPSFNQLFGLSDPV··
PTC1 31 FRRTMEDVH·····YVKNFASRLDW
P2CB 32 WRVEMEDAHTAVVGIP·····HGLEDW
CeF2 32 WRVNQEDAHNCVV·····DLHTDW
KAPP 319 RKLPMEDVCH·····YKWPLPGAN··
Conserved Motif II
e2105 stop
FEM2 196 SVLAVFDGHGGHECSQYAAGHLWET
ABI 171 HFFGVYDGHGGSQVANYCRERMHLA
HORF 193 AYFAVFDGHGGVDAARYAAVHVHTN
PTC1 52 GYFAVFDGHAGIQASKWCGKHLHTI
P2CB 54 SFFAVYDGHAGSRVANYCSTHLLEH
CeF4 52 HMFGVYDGHGGTEVSKFTSAKLPDF
KAPP 340 GLFCVCDGHGGSGAAQ·····
Conserved Motif III
abi1 D
FEM2 261 GGSTAVCCAIDMDQKLMALAWLGDSPGYVMSNIEFRQLTRGHSPSDEREARRVEEAGGQLF
ABI1 241 VGSTSVVAVVFPSHIFVANC··GDSRAVLCRGKTALPLSVDHKPDREDEAARIEAAGGKVI
HORF 254 SGTTGVCALIAGATLHVA··WLGDSQVILVQQGQVVKLMEPHRPERQDEKARIEALGGFVS
PTC1 128 SVSDDSMDLAQHQRKLYTANV·GDSRIVLFRNGNSIRLTYDHKASDTLEMQRVEQAGGLIM
P2CB 129 SGSTAVGVMISPTHIYFINC··GDSRAVLCRNGQVCFSTQDHKPCNPMEKERIQNAGGSVM
CeF4 315 SGTTACVCLVGKDKVIVANA··GDSRAVLCRNGKAVDLSVDHKPEDEVETNRIHAAGGQIE
KAPP 406 GCTATVLLVWKDNEENFFAQ··GDS·ACVIHLGG··QMTEDHRVVSLSERKRFQEAGLALR
Conserved Motif IV
FEM2 329 RVNGVLNLTRALGDVPGRPMISNEPETCQVPIESSDYL
ABI1 305 RVFGVLAMSRSIGDRYLKPSIIPDPEVTAVKRVKEDDC
HORF 319 RVNGTLAVSRAIGDVFQKPYVSGEADAASRALTGSEDY
PTC1 190 RVNGMLAVTRSLGDKFFDSLVVGSPFTTSVEITSEDKF
P2CB 190 RVNGSLAVSRALGDYDYKCVDGKGPTEQLVSPEPEVYE
CeF4 376 RVNGGLNLSRAFGDHAYKKNQELGLKEQMITALPDVKI
KAPP 473 RLFG··NLARMLGDKFPK··IS·EP···LRIDQSSKD·
Conserved Motif V
FEM2 367 VLLACDGISDVFNERDLYQLVE
ABI1 343 LILASDGVWDVMTDEEACEMAR
HORF 356 LLLACDGFFDVVPHQEVVGLVQ
PTC1 227 LILACDGLWDVIDDQDACELIK
P2CB 237 VVLACDGIWDVMSNEELCEFVN
CeF4 423 IVVACDGIWNSMESQQVVDFVR
KAPP 516 AVLASDGLWDVVSPKKAVQLVL
b245 R
Conserved Motif VI
FEM2 415 DNVSVVIG·FLRPPQDVWKLMKH
ABI1 414 DNISVVVV·DLKPRRKLKSKPLN
HORF 405 DNITVMVV·FLRDPQELLEGGNQ
PTC1 272 DNVTVMVV·FL
P2CB 286 DNMSIVLVCFANAPKVSDEAVK
CeF4 477 DNMTVICTTFDRKSKFSLKNTK
KAPP 568 DNTSIIYLDFDTSL

Figure 9. Conserved motifs alignment.

The putative FEM-2 homologues were aligned to show conserved regions. Horizontal open boxes represent the protein sequences, shaded vertical boxes represent the conserved motifs, and stippled boxes represent areas that are unique to that class of proteins. The *C. elegans* predicted F42G9.1 ORF contains an extra in-frame sequence which is indicated between motif II and III. Motif IV and V are contiguous in Class A and B proteins but are separated by a nine to ten amino acid peptide in Class C proteins. See text for details. The only protein that is not mentioned in Figure 8 is the *Schizosaccharomyces pombe ptc1*$^{+}$ (Shiozaki *et al.*, 1994) which shows the highest sequence similarity with the *Saccharomyces cerevisiae* PTC1.

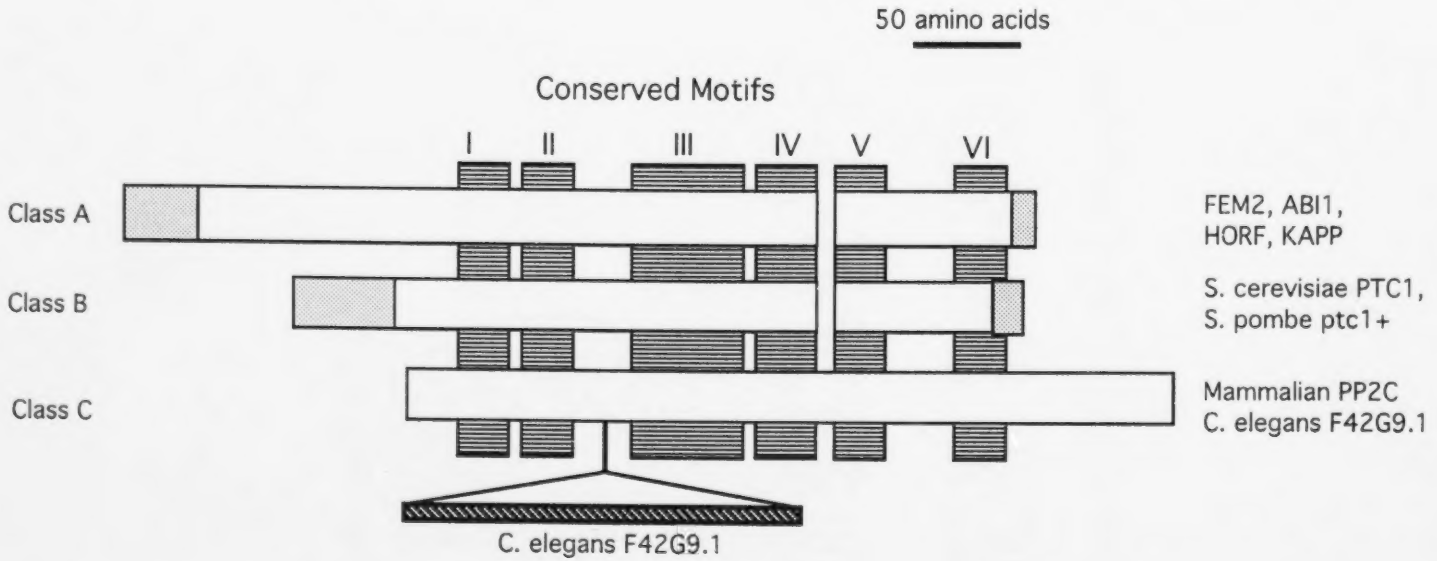
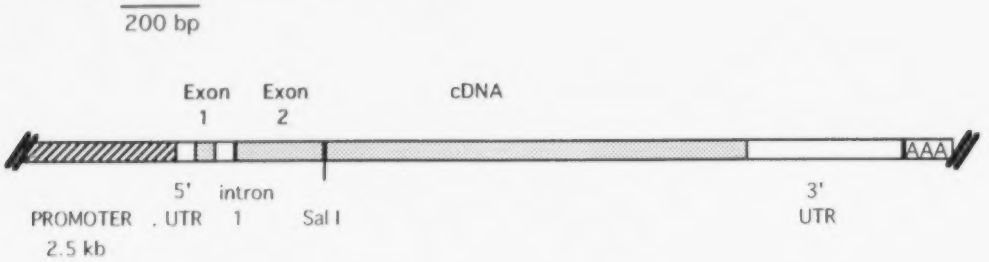

50 amino acids
Conserved Motifs
I
II
III
IV
V
VI
Class A
Class B
Class C
FEM2, ABI1,
HORF, KAPP
S. cerevisiae PTC1,
S. pombe ptc1+
Mammalian PP2C
C. elegans F42G9.1
C. elegans F42G9.1

Figure 10. *fem-2* minigene insert.

The minigene cassette was constructed by ligating a 1.4 kbp fragment of the cDNA in both orientations into a plasmid containing 2.5 kbp of upstream promoter region (pDP#AM001(+) is the proper orientation for translation and pDP#AM001(-) is in the opposite orientation). The hatched boxes represent promoter sequences, the stippled boxes represent the cDNA and coding regions, and the open boxes represent untranslated regions or introns.

pDP#AM001+



pDP#AM001-

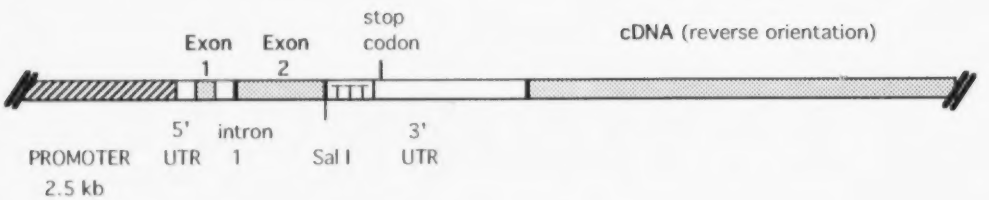


Figure 11. Rescue of *fem-2(e2105)* XO phenotype

A transgenic line of *e2105; him-8* (high incidence of males) worms carrying the minigene in the proper orientation (strain DP109 edEx36 from Table 1) was examined at 25°C and shown to rescue the male germ line and somatic phenotype (see text for details). Panel A shows a wild type hermaphrodite. *e2105* homozygous worms are transformed into fertile females at 25°C (Hodgkin, 1986) and appear identical to wild type hermaphrodites except they do not produce sperm in their germ line. Panel B shows a rescued transgenic *e2105* XO animal (strain DP109 edEx36) raised at 25°C with the arrow indicating the presence of sperm in the germ line. Panel C is a magnification of the tail of these animals showing the specialized mating structures [fan, rays (indicated by the arrow) and spike]. Panel D and E show the same structures as B and C but from a wild type XO animal to demonstrate that they are identical. In panel D the arrow indicates the presence of sperm in the gonad of the animal and the arrow in panel E points out the rays (specialized mating structure) in the male tail.

A

B

C

D

E

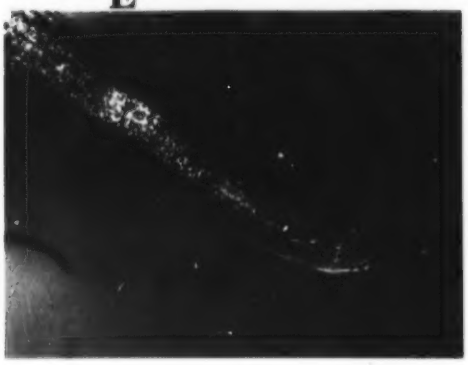


Figure 12. PCR amplification of transgenic worms

Worms were tested for the presence of the pDP#AM001(+) and pDP#AM001(-) plasmids using PCR with a primer located in the polylinker of pBluescript and one in the *fem-2* coding region as described in Results section. Marker sizes are indicated on the left in kbp. Lane 1 shows single worm amplification of a AM001(+) transgenic worm which produces a PCR product of 1.7 kbp using the primers AMC3 and T7. Lane 2 shows single worm amplification of a AM001(-) transgenic worm which also produces a PCR product of 1.7 kbp. This band represents amplification between the primers AMC3 and T7 (located at the 3' end of both constructs in the polylinker). Lane 3 is DNA from a AM001(-) transgenic worm amplified with T7 primer only. This was done to show that in the tandem array, two plasmids were not located in opposite orientations such that the T7 primer would amplify a PCR product of 1.7 kbp. Lane 4 demonstrates that worm genomic DNA from a non-transformed strain does not amplify a fragment with AMC3 and T7 primers.

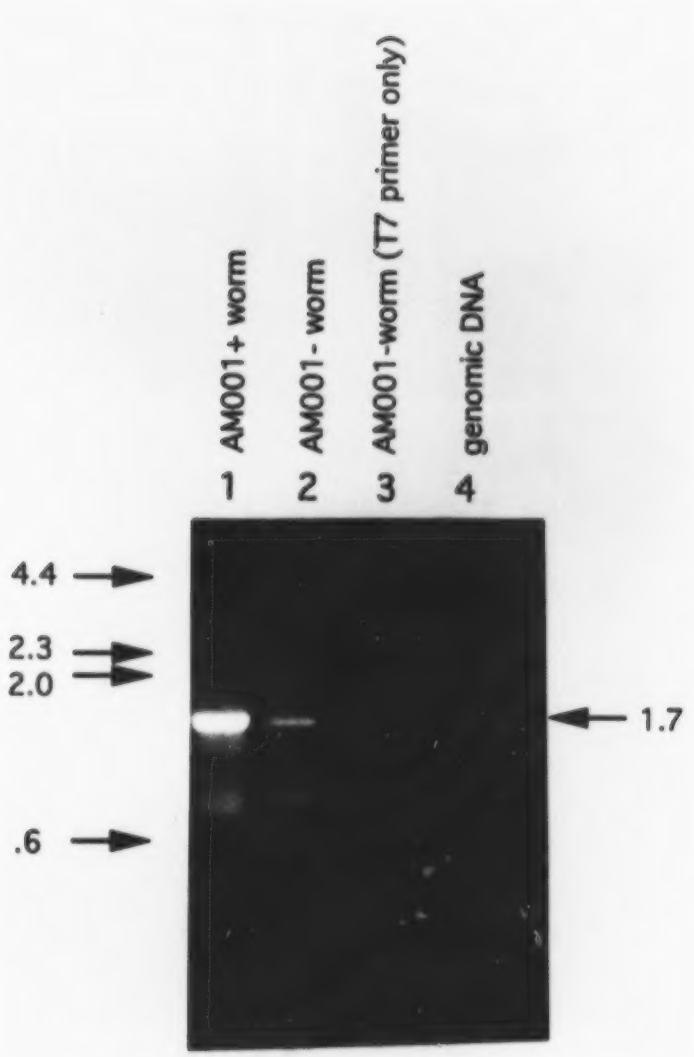
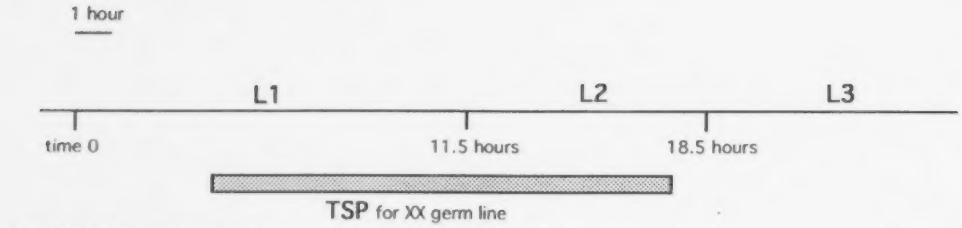

1 AM001+ worm
2 AM001- worm
3 AM001-worm (T7 primer only)
4 genomic DNA
4.4
2.3
2.0
.6
1.7

Figure 13. Heat shock results.

Worms were heat shocked as described in Methods. The developmental time line is shown with the temperature sensitive period (TSP) for *fem-2* activity in the XX germ line shown by the shaded box. L1, L2, and L3 refer to the larval stages of development. Time 0 is when the eggs hatch. Solid lines indicate duration of heat shock. Of the 48 worms carrying the 'sense' construct (pDP#AM002), none were rescued by the induced expression of *fem-2*. Of the 105 worms carrying the 'antisense' construct (pDP#AM003), a total of 17 were feminized.



pDP#AM002 trangenic line

#of worms	
4	———
4	———
14	———
20	———
6	— — — — — —

pDP#AM003 transgenic line

#of worms	
38	———
15	———
14*	———
38**	———

* 6 of these worms were feminized

** 11 of these worms were feminized

Figure 14. Suppressor screen.

The genotypes of the parental strain as well as the F1 - F3 progeny are shown. A putative mutation that suppresses *fem-2* is indicated by "sup". The *sC1* marker is a cross-over suppressor that is marked in this screen with a lethal mutation, therefore, worms that are homozygous for *sC1* are dead. See text for detailed explanation.

$\frac{\pm}{+}$; $\frac{\text{fem-2(e2105) daf-7(e1372)}}{\text{sC1(lethal)}}$ ♂

(EMS mutagenized)
where + = putative suppressor gene

↓ self

F1
at 25°C

$\frac{\text{sup}}{+}$; $\frac{\text{fem-2 daf-7}}{\text{sC1}}$	"wild type"
or	
$\frac{\text{sup}}{+}$; $\frac{\text{fem-2 daf-7}}{\text{fem-2 daf-7}}$	"dauers"
or	
$\frac{\text{sup}}{+}$; $\frac{\text{sC1(lethal)}}{\text{sC1(lethal)}}$	"dead"

where sup = new suppressor mutation

Self wild type F1 at 25°C

F2

$\frac{\text{sup}}{+}$ or $\frac{\text{sup}}{\text{sup}}$ or $\frac{\pm}{+}$	;	$\frac{\text{fem-2 daf-7}}{\text{sC1}}$ "wild type"
		or
		$\frac{\text{fem-2 daf-7}}{\text{fem-2 daf-7}}$ "dauer"
		or
		$\frac{\text{sC1(lethal)}}{\text{sC1(lethal)}}$ "dead"

Select dauers, recover at 15°C and self

F3

$\frac{\text{sup}}{+}$; $\frac{\text{fem-2 daf-7}}{\text{fem-2 daf-7}}$	"sterile"
or	
$\frac{\text{sup}}{\text{sup}}$; $\frac{\text{fem-2 daf-7}}{\text{fem-2 daf-7}}$	"fertile"
or	
$\frac{\pm}{+}$; $\frac{\text{fem-2 daf-7}}{\text{fem-2 daf-7}}$	"sterile"

Figure 15. Use of PCR to detect a Tc1 insertion event in the *fem-2* gene.

See text for the description of the screen for a Tc1 insertion mutant. In wild type DNA there is no Tc1 element in the *fem-2* region and no fragment will be amplified using the primers shown. If a Tc1 element has inserted into the *fem-2* region, a combination of primer 1 and the Tc1-specific primer or primer 2 and the Tc1-specific primer will amplify the corresponding fragment. In the experiments that were done primer 1 corresponds to DHA3 (Figure 5) and primer 2 corresponds to AMC2 (Figure 5).

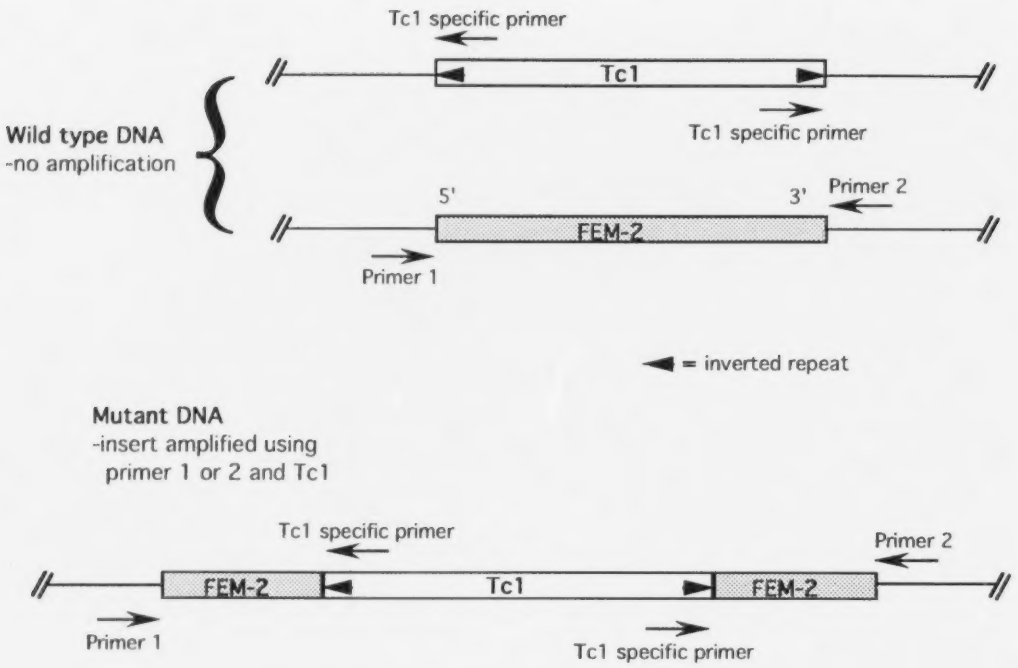

Tc1 specific primer
Tc1
Tc1 specific primer
Wild type DNA
-no amplification
5'
3'
Primer 2
FEM-2
Primer 1
= inverted repeat
Mutant DNA
-insert amplified using
primer 1 or 2 and Tc1
Tc1 specific primer
Primer 2
FEM-2
Tc1
FEM-2
Primer 1
Tc1 specific primer

Figure 16. Intronic repeat

This imperfect inverted repeat sequence was found in the second intron of the *fem-2* sequence 13 times. The numbers in brackets refers to the 3' endpoint nucleotide number of the repeat (see Figure 4). These repeats are contiguous meaning that there are no base pairs between each repeat.

```
1(624)                                  CAGCTGAAAACAGT
2(660)          GACCCACAGAGTGAACTAGCGGTGTCTTTAACTGCT
3(692)             CTTTAAGTCACTTCAACGCAAGTCTCTACAAGT
4(728)          CATATTAGGACTATAAATACAATGGCTATCATGGT
5(749)               GTCCATTAATAATTTACACGT
6(779)               TTTAGACGATAATTAAATCGTCTGGTTCGT
7(813)             ACCCGGACGATTTATAAGTCGTCCGGATAGGGGT
8(843)             AACCAGACGATTTATAAATCGTATGGATTG
9(877)            AACCCAGACGATTTATAAATCGTCTGGTAGGGGT
10(908)            AACCAGACGATTAATAAATCGTCTGGTTGAT
11(939)            ACCCAGACGATTAATAAATCGTCTGGTGGAT
12(971)            ACCCAGACGATTTATAAATCGTATGGATGGT
13(998)            ACTCAGACGATTTATAAATCGTCTGGAT
```

Figure 17. Model for FEM-2 activity.

FEM-2 may act on FEM-1 by dephosphorylating it and allowing it to inactivate TRA-1. This would result in the worm developing a male soma. If FEM-2 does not dephosphorylate FEM-1, it is unable to bind and inactivate TRA-1, therefore allowing the development of a female soma. See text for details.

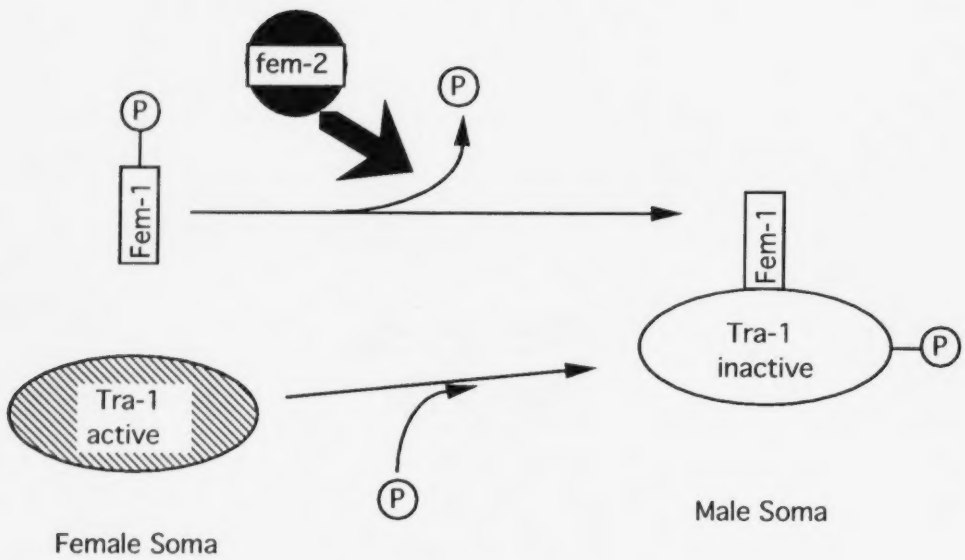

fem-2
P
P
Fem-1
Fem-1
Tra-1
inactive
P
Tra-1
active
P
Female Soma
Male Soma

Figure 18. 5' ends of the cDNA clones

The genomic DNA that corresponds to the 5' end of the *fem-2* cDNA is shown at the top and the three cDNA clones that contained a partial SL1 sequence are aligned below. The complete sequence of SL1, as well as the portions present on the cDNA clones is shown (underlined). The lower case letters refer to genomic sequence and capital letters refer to cDNA sequence or partial SL1 sequences.

```
genomic          ...aaaaacaattattgaaatgtcagATCAACCCAGCTGTCCC...
pDP#AMc8                                GAGATCAACCCAGCTGTCCC...
pDP#AMc12                     CCAAGTTTGAGATCAACCCAGCTGTCCC...
SL1                 GGTTTAATTACCCAAGTTTGAG
pDPAMc3           ...AGGGCGGCCAATCGAGTTTGAGATCAACCCAGCTGTCCC...
                     (  Chimaeric  )
```

Bibliography

Ahringer, J., and Kimble, J. 1991. Control of the sperm-oocyte switch in *Caenorhabditis elegans* hermaphrodites by the *fem-3* 3' untranslated region. *Nature* 349: 346-348.

Ahringer, J., Rosenquist, T., Lawson, D., and Kimble, J. 1992. The *Caenorhabditis elegans* sex determining gene *fem-3* is regulated post-transcriptionally. *EMBO J.* 11: 2303-2310.

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Barstead, B. and Waterston, R. 1989. The basal component of the nematode dense-body is vinculin. *J. Biol. Chem.* 264: 10177-10185.

Barton, M. and Kimble, J. 1990. *fog-1*, a regulatory gene required for specification of spermatogenesis in the germline of *Caenorhabditis elegans*. *Genetics* 125: 29-39.

Barton, M., Schedl, T., and Kimble, J. 1987. Gain-of-function mutations of *fem-3*, a sex determination gene in *C. elegans*. *Genetics* 115: 107-119.

Berg, D., and Howe, M. (eds.). 1989. *Mobile DNA*. American Society for Microbiology, Washington, D.C.

Brenner, S. 1974. The genetics of *C. elegans*. *Genetics* 77: 71-94.

Cohen, P. 1989. The structure and regulation of protein phosphatases. *Ann. Rev. Biochem.* 58: 453-508.

Collins, J., Saari, B., and Anderson, P. 1987. Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature (London)* 328: 726-728.

Collins, J., Forbes, E., and Anderson, P. 1989. The Tc3 family of transposable genetic elements in *Caenorhabditis elegans*. *Genetics* 121: 47-55.

Coulson, A., Waterston, R., Kiff, J., Sulston, J., and Kohara, Y. 1988. Genome linking with yeast artificial chromosomes. *Nature* 335: 184-186.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulson, J., and Waterston, R. 1991. YACs and the *C. elegans* genome. *Bioessays* 13: 413-417.

deBono, M., Zarkower, D., Hodgkin, J. 1995. Dominant feminizing mutations implicate protein-protein interactions as the main mode of regulation of the nematode sex-determining gene *tra-1*. *Genes & Dev.* 9: 155-167.

DeLong, L., Plenefisch, J., Klein, R., and Meyer, B. 1993. Feedback control of sex determination by dosage compensation revealed through *Caenorhabditis elegans sdc-3* mutations. *Genetics* 133: 875-896.

Doniach, T. 1986. Activity of the sex-determining gene Tra-2 is modulated to allow spermatogenesis in the *C. elegans* hermaphrodite. *Genetics* 114: 53-76.

Doniach, T. and Hodgkin, J. 1984. A sex-determining gene, *fem-1*, required for both male and hermaphrodite development in *Caenorhabditis elegans*. *Dev. Biol.* 106: 223-235.

Ellis, R. and Kimble, J. 1995. The *fog-3* gene and regulation of cell fate in the germ line of *Caenorhabditis elegans*. *Genetics* 139: 561-577.

Emmons, S. 1988. "The Genome," In The Nematode *Caenorhabditis elegans* (ed. W.B.Wood), pp. 47-81. Cold Spring Harbor Laboratory, New York.

Epstein, H., and Thomas, J. 1974. Temperature sensitive mutation affecting myofilament assembly in *C. elegans*. *Nature* 250: 579-580.

Evans, T., Goodwin, E., and Kimble, J. 1992. Translational regulation of development and maternal RNAs in *Caenorhabditis elegans*. *Sem. Dev. Biol.* 3: 381-389.

Files, J., Carr, S., and Hirsh, D. 1983. Actin gene family of *Caenorhabditis elegans*. *J. Mol. Biol.* 164: 355-375.

Fire, A. 1986. Integrative transformation of *Caenorhabditis elegans*. *EMBO J.* 10: 2673-2680.

Frohman, M., Dush, M., and Martin, G. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Nat. Acad. Sci. U.S.A.* 85: 8998-9002.

Goodwin, E., Okkema, P., Evans, T., and Kimble, J. 1993. Translational regulation of *tra-2* by its 3' untranslated region controls sexual identity in *C. elegans*. *Cell* 75: 329-339.

Gough, J. and Murray, N. 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. *J. Mol. Biol.* 166: 1-19.

Graham, P. and Kimble, J. 1993. The *mog-1* gene is required for the switch from spermatogenesis to oogenesis in *Caenorhabditis elegans*. *Genetics* 133: 919-931.

Graham, P., Schedl, T., and Kimble, J. 1993. More *mog* genes that influence the switch from spermatogenesis to oogenesis in the hermaphrodite germ line of *Caenorhabditis elegans*. *Developmental Genetics* 14: 471-484.

Hengartner, M. and Horvitz, R. 1994. Programmed cell death in *Caenorhabditis elegans*. *Curr. Opin. Gen. Dev.* 4:581-586.

High, S. and Dobberstein, B. 1992. Mechanisms that determine the transmembrane disposition of proteins. *Curr. Opin. Cell Biol.* 4: 581-586.

Hirsh, D., Oppenheim, D., and Klass, M. 1976. Development of the reproductive system of *Caenorhabditis elegans*. *Dev. Biol.* 49: 200-219.

Hirsh, D., Emmons, S., Files, J., and Klass, M. 1979. Stability of the *Caenorhabditis elegans* gemome during development and evolution. *ICN-UCLA Symp. Mol. Cell. Biol.* 14: 205-218.

Hodgkin, J. 1980. More sex determination mutants of *Caenorhabditis elegans*. *Genetics* 96: 649-664.

Hodgkin, J. 1983. Two types of sex determination in a nematode. *Nature* 304: 267-268.

Hodgkin, J. 1983. X chromosome dosage and gene expressionin *Caenorhabditis elegans:* two unusual dumpy genes. *Mol. Gen. Genet.* 192: 452-458.

Hodgkin, J. 1985. Males, hermaphrodites and females: Sex determination in *C. elegans*. *Trends in Genetics* 1: 85-88.

Hodgkin, J. 1986. Sex determination in the nematode *C. elegans*: analysis of *tra-3* suppressors and characterization of *fem* genes. *Genetics* 114: 15-52.

Hodgkin, J. 1987. A genetic analysis of the sex determining gene, *tra-1*, in the nematode *C. elegans*. *Genes Dev.* 1: 731-745.

Hodgkin, J. 1987. Primary sex determination in the nematode *Caenorhabditis elegans*. *Development* 101(Suppl.): 5-16.

Hodgkin, J. 1987. Sex determination and dosage compensation in *C. elegans*. *Annual Review of Genetics* 21: 133-154.

Hodgkin, J. 1988. "Sexual dimorphism and sex determination," In The Nematode *Caenorhabditis elegans*. (ed. W.B. Wood), pp. 243-281. Cold Spring Harbor Laboratory, New York.

Hodgkin, J. 1990. Sex determination compared in *Drosophila* and *Caenorhabditis*. *Nature* 344: 721-728.

Hodgkin, J. 1992. Genetic sex determination mechanisms and evolution. *Bioessays* 14: 253-261.

Hodgkin, J. 1993. Molecular cloning and duplication of the nematode sex determining gene *tra-1*. *Genetics* 133: 543-560.

Hodgkin, J. and Brenner, S. 1977. Mutations causing transformation of sexual phenotype in the nematode *C. elegans*. *Genetics* 86: 275-287.

Hodgkin, J., Chisholm, A., and Shen, M. 1989. Major sex-determining genes and the control of sexual dimorphism in *Caenorhabditis elegans*. *Genome* 31: 625-637.

Hodgkin, J., Doniach, T., and Shen, M. 1985. Sex determination pathway in the nematode *C. elegans*: Variations on a theme. *Cold Spring Harbor Symposia on Quantitative Biology*, Vol 50: 585-594.

Hodgkin, J., Horvitz, H., and Brenner, S. 1979. Non disjunction mutants of the nematode *C. elegans*. *Genetics* 91: 67-94.

Hunter, C. and Wood, W. 1990. The *tra-1* gene determines sexual phenotype cell-autonomously in *C. elegans*. *Cell* 63: 1193-1204.

Hunter, C. and Wood, W. 1992. Evidence from mosiac analysis of the masculinizing gene *her-1* for cell interactions in *C. elegans* sex determination. *Nature* 355: 551-555.

Johnson, T. 1995. Analysis of the expression patterns of the *fem-2* gene of *Caenorhabditis elegans*. M. Sc. thesis, University of Alberta

Kataoka, T., Broek, D., and Wigler, M. 1985. DNA sequence and characterization of the *S. cerevisiae* gene encoding adenylate cyclase. *Cell* 43: 493-505.

Kayne, P. 1994. Knocking down a pathway to build it up. *Trends in Genetics* 10: 145-147.

Kenyon, C. 1985. Cell lineage and the control of *Caenorhabditis elegans* development. *Philos. Trans. R. Soc. Lond. Biol.* 312: 21-38.

Kimble, J., Edgar, L., and Hirsh, D. 1984. Specification of male development in *Caenorhabditis elegans*: The *fem* Genes. *Dev. Biol.* 105: 234-239.

Kimble, J. and Hirsch, D. 1979. The postembryonic lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev. Biol.* 70: 396-417.

Klass, M., Wolf, N., and Hirsch, D. 1976. Development of the male reproductive system and sexual transformation in the nematode *C. elegans*. *Dev. Biol.* 52: 1-18.

Kleckner, N. 1981. Transposable elements in prokaryotes. *Ann. Rev. Genet.* 15: 341-404.

Kramer, J., French, R., Park, E., and Johnson, J. 1990. The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. *Mol. Cell. Biol.* 10: 2081-2089.

Krause, M. and Hirsh, D. 1987. Actin gene expression in *Caenorhabditis elegans*. In *Molecular biology of the cytoskeleton* (ed. G.Borisy *et al.*), pp.287-292. Cold Spring Harbor Laboratory, New York.

Kuwabara, P., and Kimble, J. 1992. Molecular genetics of sex determination in *C. elegans*. *Trends in Genetics* 8: 164-168.

Kuwabara, P., Okkema, P., and Kimble, J. 1992. *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway. *Mol. Biol. Cell.* 3: 461-473.

Lackner, M., Kornfield, K., Miller L., Horvitz, H., and Kim, S. 1994. A MAP kinase homologue, *mpk-1*, is involved in ras-mediated induction of vulval fates in *Caenorhabditis elegans*. *Genes Dev.* 8: 163-173.

Leung, J., Bouvier-Durand, M., Morris, P., Guerrier, D., Chefdor, F., and Giraudat, J. 1994. Arabidopsis ABA-response gene ABI1: features of a calcium-modulated protein phosphatase. *Science* 264: 1448-1452.

Lux, S., John, K., and Bennett, V. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homologuey to tissue-differentiation and cell-cycle control proteins. *Nature* 344: 36-42.

Madl, J. and Herman, R. 1979. Polyploids and sex determination in *Caenorhabditis elegans*. *Genetics* 93: 393-402.

Maeda, T., Tsai, A., and Saito, H. 1993. Mutations in a protein tyrosine phosphatase gene (PTP2) and a protein serine/threonine phosphatase gene (PTC1) cause a synthetic growth defect in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 13: 5408-5417.

Maeda, T., Wurgler-Murphy, S., and Saito, H. 1994. A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* 369: 242-245.

Meester, F., Bracha, R., Huber, M., Keren, Z., Rozenblatt, S., and Mirelman, D. 1991. Cloning and characterization of an unusual elongation factor-1 cDNA from *Entamoeba histolytica*. *Mol. Biochem. Parasit.* 44: 23-32.

Mello, C., Kramer, J., Stinchcomb, D., and Ambros, V. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* 10: 3959-3970.

Meneely, P. 1990. X-linked gene expression and sex determination in *Caenorhabditis elegans*. *Bioessays* 12: 513-517.

Meneely, P., and Herman, R. 1979. Lethals, steriles and deficiencies in a region of the X chromosome of *C. elegans*. *Genetics* 92: 99 -115.

Meyer, K., Leube, M., and Grill, E. 1994. A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. *Science* 264: 1452-1455.

Miller, L., Plenefisch, J., Casson, L., and Meyer, B. 1988. *xol-1*: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in *C. elegans*. *Cell* 55: 167-183.

Nelson, G., Lew, K., and Ward, S. 1978. Intersex, a temperature sensitive mutant of the nematode *C. elegans*. *Dev. Biol.* 66: 386-409.

Nomarski, G. 1955. Microinterféromètre differentiel á ondes polarisées. *J. Phys. Radium* 16: 9-13.

Nomura, N., Miyajima, N., Kawarabayashi, Y., and Tabata, S. 1993. Prediction of new human genes by entire sequencing of randomly sampled cDNA clones. documentation from Genbank accession #D13640.

Nusbaum, C. and Meyer, B. 1989. The *Caenorhabditis elegans* gene *sdc-2* controls sex determination and dosage compensation in XX animals. *Genetics* 122: 579-593.

Okkema, P. and Kimble, J. 1991. Molecular analysis of *tra-2*, a sex determining gene in *C. elegans. EMBO J.* 10: 171-176.

Pallen, C., Tan, Y., and Guy, G. 1992. Protein phosphatases in cell signalling. *Curr. Opin. Cell Biol.* 4: 1000-1007.

Parkhurst, S. and Meneely, P. 1994. Sex determination and dosage compensation: Lessons from flies and worms. *Science* 264: 924-932.

Perry, M., Li, W., Trent, C., Robertson, B., Fire, A., Hageman, J., and Wood, W. 1993. Molecular characterization of the *her-1* gene suggests a direct role in cell signaling during *Caenorhabditis elegans* sex determination. *Genes and Development* 7: 216-228.

Pilgrim, D. 1993. The genetic and RFLP characterization of the left end of linkage group III in *Caenorhabditis elegans. Genome* 36: 712-724.

Pilgrim, D. and Bell, J. 1993. Expression of a *Drosphila melanogaster* amber suppressor tRNASer in *Caenorhabditis elegans. Mol. Gen. Genet.* 241: 26-32.

Pilgrim, D., McGregor, A., Jackle-Baldwin, P., Johnson, T., and Hansen, D. 1995. The *C. elegans* sex-determining gene *fem-2* encodes a putative protein phosphatase. *Mol. Biol. Cell* (in press).

Plenefisch, J., DeLong, L., and Meyer, B. 1989. Genes that implement the hermaphrodite mode of dosage compensation in *Caenorhabditis elegans. Genetics* 121: 57-76.

Riddle, D. 1988. "The dauer larvae", in The Nematode *Caenorhabditis elegans*. (ed. W.B. Wood), pp. 393-415. Cold Spring Harbor Laboroatory, New York.

Rosenquist, T. and Kimble, J. 1988. Molecular cloning and transcript analysis of *fem-3*, a sex determining gene in *Caenorhabditis elegans. Genes Dev.* 2: 606-616.

Rosenzweig, B., Liao, L., and Hirsh, D. 1983. Sequence of the *C. elegans* transposable element Tc1. *Nucleic Acids Res.* 11: 7137-7140.

Rushforth, A., Saari, B., and Anderson, P. 1993. Site-selected insertion of the transposable Tc1 into a *Caenorhabditis elegans* myosin light chain gene. *Mol. Cell. Biol.* 13: 902-910.

Sambrook, J., Fritsch, E., and Maniatis, T. 1989. "Molecular Cloning, a laboratory manual", 2nd. edition. Cold Spring Harbor Laboratory Press, N.Y.

Sanger, F., Nicklen, S., and Coulson, A. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.

Schedin, P., Jonas, P., and Wood, W. 1994. Function of the *her-1* gene is required for maintenance of male differentiation in adult tissues of *C. elegans*. *Dev. Gen.* 15: 201-204.

Schedl, T. and Kimble, J. 1988. *fog-2*, a germ-line specific sex determination gene required for hermaphrodite spermatogenesis in *Caenorhabditis elegans*. *Genetics* 119: 43-61.

Schedl, T., Graham, P., Barton, M., and Kimble, J. 1989. Analysis of the role of *tra-1* in germline sex determination in the nematode *Caenorhabditis elegans*. *Genetics* 123: 755-769.

Schneider, T., Stormo, G., Gold, L., and Ehenfeucht, A. 1986. Information content of binding sites on nucleotide sequences. *J. Mol. Biol.* 188: 415-431.

Selfors, L. and Stern, M. 1994. MAP kinase function in *C. elegans*. *Bioessays* 16: 301-304.

Sheets, M., Ogg, S., and Wickens, M. 1990. Point mutations in AAUAAA and the poly(A) site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucl. Acids Res.* 18: 5799-5805.

Shen, M. and Hodgkin, J. 1988. *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* 54: 1019-1031.

Shiozaki, K., Akhavan-Niaki, H., McGowan, C., and Russel, P. 1994. Protein phosphatase 2C, encoded by *ptc1*$^+$, is important in the heat shock response of *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* 14: 3742-3751.

Siddiqui, S. 1990. Mutations affecting axonal growth and guidance of motor neurons and mechanosensory neurons in the nematode *Caenorhabditis elegans*. *Neurosci. Res.* 13: S171-S190.

Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. 1993. Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell* 73: 521-532.

Spence, A., Coulson, A., and Hodgkin, J. 1990. The product of *fem-1*, a nematode sex-determining gene, contains a motif found in cell cycle control proteins and receptors for cell-cell interactions. *Cell* 60: 981-990.

Stone, J., Collinge, M., Smith, R., Horn, M., and Walker, J. 1994. Interaction of a protein phosphatase with an *Arabidopsis* serine-threonine receptor kinase. *Science* 266: 793-795.

Stringham, E., Dixon, D., Jones, D., and Candido, E. 1992. Temporal and spatial expression patterns of the small heat shock (HSP16) genes in transgenic *Caenorhabditis elegans. Mol.Biol. Cell* 3: 221-233.

Sulston, J. and Brenner, S. 1974. The DNA of *Caenorhabditis elegans. Genetics* 77: 95-104.

Sulston, J. and Horvitz, H. 1977. Post-embryonic cell lineages of the nematode *Caenorhabditis elegans. Dev. Biol.* 56: 110-156.

Sulston, J., Schierenberg, E., White, J., and Thomson, A. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans. Dev. Biol.* 100: 64-119.

Sulston, J. and Hodgkin, J. 1988. "Methods, " In the Nematode *Caenorhabditis elegans.* (ed. W.B. Wood), pp. 587-606. Cold Spring Harbor Laboratory, New York.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qui, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., and Waterston, R. 1992. The *C. elegans* genome sequencing project: a beginning. *Nature* 356: 37-41.

Trent, C., Wood, W., and Horvitz, H. 1988. A novel dominant transformer allele of the sex-determining gene *her-1* of *C. elegans. Genetics* 120: 145-157.

Trent, C., Purnell, B., Gavinski, S., Hagemen, J., Chamblin, C., and Wood, W. 1991. Sex-specific transcriptional regulation of the *C. elegans* sex-determining gene *her-1. Mech. Dev.* 34: 43-56.

Villeneuve, A. and Meyer, B. 1990. The role of *sdc-1* in the sex determination and dosage compensation decisions in *Caenorhabditis elegans. Genetics* 124: 91-114.

Wang, Y. and Roach, P. 1993. Inactivation of rabbit muscle glycogen synthase by glycogen synthase kinase-3. *J.Biol. Chem.* 268: 23876-23880.

Waterston, R. 1981. A second informational suppressor, *sup-7* X, in *C. elegans. Genetics* 97: 307-327.

Waterston, R., Thomson, J., and Brenner, S. 1980. Mutants with altered muscle structure in *C. elegans. Dev. Biol.* 77: 271-302.

Wickens, M. 1990. In the beginning is the end: regulation of poly (A) addition and removal during early development. *Trends Biochem. Sci.* 15: 320-324.

Wenk, J., Trompeter, H., Pettrich, K., Cohen, P., Campbell, D., and Mieskes, G. 1992. Molecular cloning and primary structure of a protein phosphatase 2C isoform. *FEBS Lett.* 297: 135-138.

Williams, B., Schrank, B., Huynh, C., Shownkeen, R., and Waterston, R. 1992. A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics* 131: 609-624.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Smaldon, N., Smith, A., Sonnhammer, E., Staden, R., Sulston, J., Terry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. 1994. The *C. elegans* genome project: Contiguous nucleotide sequence of over two megabases from chromosome III. *Nature* 368: 32-38.

Wood, W. 1988. "Introduction to *C. elegans* biology, " In The Nematode *Caenorhabditis elegans* . (ed. W.B. Wood), pp. 1-17. Cold Spring Harbor Laboroatory, New York.

Wood, W. 1993. Molecular characterization of the *her-1* gene suggests a direct role in cell signalling during *Caenorhabditis elegans* sex determination. *Genes Dev.* 7: 216-228.

Zarkower, D. and Hodgkin, J. 1992. Molecular analysis of the *C. elegans* sex-determining gene *tra-1*: a gene encoding two zinc finger proteins. *Cell* 70: 237-249.

Zarkower, D. and Hodgkin, J. 1993. Zinc fingers in sex determination: only one of the two *C. elegans* Tra-1 proteins bind DNA *in vitro. Nucleic Acids Res.* 21: 3691-3698.

Zarkower, D., DeBono, M., Aronoff, R., and Hodgkin, J. 1994. Regulatory rearrangements and smg-sensitive alleles of the *C. elegans* sex-determining gene *tra-1. Dev. Gen.* 15: 240-250.

Appendix

M9 buffer

KH_2PO_4	3g
Na_2HPO_4	6g
NaCl	5g
$MgSO_4$ 1M	1ml

The solution was made up to 1 liter with distilled water and sterilized by autoclaving for 20 minutes at 121°C.

Freezing Solution

NaCl	5.85g
KH_2PO_4	6.8g
glycerol	300g
1M NaOH	5.6ml

Solution was made up to 1 liter with distilled water and sterilized by autoclaving for 20 minutes at 121°C. Then, using sterile technique, add 3ml of $MgSO_4$ 1M.

EMS solution

M9 buffer	2ml
liquid EMS	0.02ml

Gently agitate until the dense oily liquid has dissolved then add 2ml of a suspension of washed worms in M9 buffer to this solution (i.e., final concentration 50mM.

Worm Lysis Buffer

KCl 1M	500ul
Tris pH 8.0 1M	70ul
$MgCl_2$ 1M	25ul
NP-40	50mg
Tween 20	50mg
gelatin 2%	50ul
H_2O	9.2ml

Filter sterilize this 10ml solution and store at room temperature. Just before use, add 6ul of 10mg/ml proteinase K (in H_2O and stored in aliquots at -20°C) to 1ml of lysis buffer.